

BIRCH'S THEOREM IN FUNCTION FIELDS

SIU-LUN ALAN LEE

ABSTRACT. We establish an asymptotic formula for the number of points with coordinates in $\mathbb{F}_q[t]$ on a complete intersection of degree d defined over $\mathbb{F}_q[t]$, with explicit error term, provided that the characteristic of \mathbb{F}_q is greater than d , the codimension of the singular locus of the complete intersection is large enough, and this intersection has a non-singular point at each place of $\mathbb{F}_q[t]$. In particular, when this complete intersection is non-singular, we show that it satisfies weak approximation.

1. INTRODUCTION

In his seminal paper [2], Birch considered the problem of counting points with integer coordinates on the complete intersection of a number of hypersurfaces defined over \mathbb{Z} . He established an asymptotic formula for the number of such points in an expanding box, under some conditions on the geometry of the complete intersection. On the basis of the close resemblance between the rings \mathbb{Z} and $\mathbb{F}_q[t]$, we consider the analogue of Birch's problem over $\mathbb{F}_q[t]$. To state our problem more precisely, let s, d, R be positive integers with $s > R$ and $d \geq 2$, and let q be a power of a prime p , so that the field \mathbb{F}_q has characteristic p . Write $\mathbf{x} = (x_1, \dots, x_s)$, and let X denote the complete intersection defined by the simultaneous equations

$$F_1(\mathbf{x}) = \dots = F_R(\mathbf{x}) = 0, \quad (1.1)$$

where F_1, \dots, F_R are forms of degree d in s variables with coefficients in $\mathbb{F}_q[t]$. The aim of this paper is to generalise Birch's result to the function field setting, as well as to establish weak approximation on X .

Before we can state our results precisely, we have to introduce some further notation. For simplicity, write $\mathbb{A} = \mathbb{F}_q[t]$ and $\mathbb{K} = \mathbb{F}_q(t)$. For any non-zero $x \in \mathbb{A}$, write $\text{ord } x$ for its degree. By convention, we define $\text{ord } 0 = -\infty$. We extend the function ord to \mathbb{K} by letting $\text{ord}(x/y) = \text{ord } x - \text{ord } y$ whenever $x, y \in \mathbb{A}$ with $y \neq 0$. We then define an absolute value $\langle \cdot \rangle : \mathbb{K} \rightarrow \mathbb{R}$ by taking

$$\langle \alpha \rangle = q^{\text{ord } \alpha} \quad \text{when } \alpha \in \mathbb{K}. \quad (1.2)$$

2000 *Mathematics Subject Classification.* 11D72, 11P55.

Key words and phrases. Birch's theorem, Hardy-Littlewood method, function fields, weak approximation.

The author is supported by the University of Bristol Overseas Centenary Postgraduate Research Scholarship.

Let \mathbb{K}_∞ denote the completion of \mathbb{K} under this absolute value. We extend this absolute value to \mathbb{K}_∞^m , for any positive integer m , by

$$\langle \alpha \rangle = \max \{ \langle \alpha_j \rangle : j = 1, \dots, m \}. \quad (1.3)$$

As a source of motivation, the sets \mathbb{A} , \mathbb{K} , \mathbb{K}_∞ are the function field analogues of \mathbb{Z} , \mathbb{Q} and \mathbb{R} respectively. With this in mind, we call any tuple with coordinates in \mathbb{A} , \mathbb{K} and \mathbb{K}_∞ an *integral point*, a *rational point* and a *real point* respectively. It is also convenient to write $\hat{Y} = q^Y$ for any real number Y . On the other hand, we need an analogue of the p -adic numbers over function fields¹. In the sequel, we reserve the symbol ϖ for monic irreducible polynomials in \mathbb{A} . For any ϖ , let \mathbb{K}_ϖ denote the completion of \mathbb{K} with respect to the place at ϖ . We refer to this completion as the ϖ -*adic completion* of \mathbb{K} , and the tuples with coordinates in \mathbb{K}_ϖ as ϖ -*adic points*. Finally let $\mathbf{F} = (F_1, \dots, F_R)$. For any s -tuple \mathbf{x} , write $\nabla \mathbf{F}(\mathbf{x})$ for the matrix $(\partial F_i / \partial x_j)$ of derivatives of the forms F_i , where $i = 1, \dots, R$ and $j = 1, \dots, s$. Let $\text{Sing } X$ denote the singular locus of X . More precisely, $\text{Sing } X$ consists of all points $\mathbf{x} \in X$, such that

$$\text{rank } \nabla \mathbf{F}(\mathbf{x}) < R.$$

Any point in $X \setminus \text{Sing } X$ is called a *non-singular point* of X . Note that our singular locus differs from Birch's singular locus in [2] in notation. But Aleksandrov and Moroz (see Corollary of [1]) have shown that the two singular loci in fact coincide.

By the Lang-Tsen theorem (see Theorem 3.6 of [5]), whenever $s > Rd^2$, there exists a non-trivial rational point on X . It is natural now to seek an asymptotic formula for the number of rational points on X . In the sequel, let P be a large positive real number that is allowed to tend to infinity. Write $\rho(P)$ for the number of integral points on X with $\langle \mathbf{x} \rangle < \hat{P}$. The asymptotic formula for this quantity is unveiled below.

Theorem 1.1. *Write*

$$k = s - \dim \text{Sing } X - R(R+1)(d-1)2^{d-1}. \quad (1.4)$$

Assume that $k \geq 1$, $p > d$, and that X contains a non-singular real point as well as a non-singular ϖ -adic point for every ϖ . Then, as $P \rightarrow \infty$, we have

$$\rho(P) = \hat{P}^{s-Rd} (\mathfrak{S} \mathcal{J} + O(\hat{P}^{-k/(2^{d+1}R(d-1))})),$$

where \mathfrak{S} , \mathcal{J} and the implied constant in the $O(\cdot)$ -notation are positive constants depending only on s, d, R, q and X .

The quantities \mathfrak{S} and \mathcal{J} are defined in §7 and §8 respectively. The former, called the *singular series*, is equal to the product of densities of ϖ -adic solutions of the equations (1.1), over all ϖ . Meanwhile, the *singular integral* \mathcal{J} is equal to the density of real solutions of the same equations. The number of variables required as asserted in Theorem 1.1 is identical to the analogous situation, where \mathbb{A} is replaced by \mathbb{Z} , treated in Theorem 1 of [2]. Aside from establishing

¹This is the only instance in the paper where p refers to an arbitrary rational prime, and has nothing to do with the characteristic of \mathbb{F}_q .

an asymptotic formula with \mathbb{A} replacing \mathbb{Z} , the major improvement here is the appearance of an explicit error term in our asymptotic formula. Now that we work over a field with positive characteristic p , our results inevitably put a condition on p . In the theorem above, the condition that p be larger than the degree d of the forms involved is due to the Weyl differencing argument employed in §3.

For any field $L \supseteq \mathbb{K}$, write $X(L) = X \cap L^s$. Recall that X satisfies *weak approximation* if the image of the canonical embedding

$$X(\mathbb{K}) \hookrightarrow X(\mathbb{K}_\infty) \times \prod_{\varpi} X(\mathbb{K}_\varpi)$$

is dense. By adapting the proof of Corollary 1 in §5 of [16] to the function field setting, we prove that the complete intersection X given in (1.1) indeed satisfies this property.

Theorem 1.2. *Let X be non-singular. Suppose $p > d$ and*

$$s \geq R(R+1)(d-1)2^{d-1}.$$

Assume further that $X(\mathbb{K}_\infty) \neq \emptyset$ and $X(\mathbb{K}_\varpi) \neq \emptyset$ for every ϖ . Then weak approximation holds on X .

As far as the author is concerned, this result does not seem to be a trivial consequence of the Lang-Tsen theorem. In Theorem 1 of [7], Hassett and Tschinkel proved that the function $\phi : \mathbb{N} \rightarrow \mathbb{N}$, given by $\phi(1) = 1$ and

$$\phi(d) = \binom{\phi(d-1) + d - 1}{\phi(d-1)} \quad (d > 1),$$

satisfies the property that, when \mathbb{K} is replaced by the function field of a smooth curve over any algebraically closed field, $R = 1$ and $s > \phi(d)$, the hypersurface X satisfies weak approximation. So when X is a quadratic (resp. cubic) hypersurface, their result requires 3 (resp. 7) variables, whereas Theorem 1.2 requires 4 (resp. 16) variables. However, as soon as $d \geq 4$, the quantity $\phi(d)$ is much larger than $R(R+1)(d-1)2^{d-1} = 2^d(d-1)$.

The plan of this paper is as follows. In §2 the foundations for the use of the Hardy-Littlewood method in the function field setting are put in place, with the view towards relating the counting functions involved to an integral of an exponential sum. In §3 the technique of Weyl differencing is employed in establishing an upper bound for our exponential sum that is sufficiently generic with respect to changes in parameters. In §4 we introduce the key ingredient of the Hardy-Littlewood method, namely the dissection of a suitable compact group into major and minor arcs. In §5 an estimate for the contribution from the minor arcs to the integral in question is obtained. In the following sections, we turn our attention towards the major arc contribution. In §6 the exponential sum of interest is expressed on the major arcs in terms of functions that are suitably well behaved, so as to streamline our subsequent analysis. In §7 and §8 respectively we investigate the singular series and singular integral. Via an application of the upper bound obtained in §3, the major arc contribution

is expressed in terms of these two quantities, allowing for a permissible error in the process. These two quantities are shown to be positive, so as to render our major arc analysis meaningful. Finally, in §9, the minor arc estimate in §5 is combined with the asymptotic formula for the major arc contribution. With different choices of parameters, Theorems 1.1 and 1.2 follow at once. Two appendices are included at the end of the paper. They provide essential components for our analysis in the main text. The first deals with the change-of-variable property for integrals over function fields. The second concerns the properties of lattices defined over function fields.

The author would like to thank his PhD supervisor Professor Trevor Wooley for his patient guidance and invaluable comments on his work. Without his help, this paper would never have been possible.

2. THE HARDY-LITTLEWOOD METHOD OVER FUNCTION FIELDS

The application of the Hardy-Littlewood method in the setting where integers are replaced by polynomials in \mathbb{A} requires some explanation. Our goal here is to introduce such notation and basic notions as are subsequently needed to initiate discussion of the key components of this version of the circle method. The material here is taken from section 1 of [10] and section 2 of [11].

By definition, every element of \mathbb{K}_∞ has the form

$$\alpha = \sum_{i=-\infty}^n a_i t^i, \quad (2.1)$$

where $n \in \mathbb{Z}$ and $a_i \in \mathbb{F}_q$ for all integers $i \leq n$. Put

$$\mathbb{T} = \{\alpha \in \mathbb{K}_\infty : \langle \alpha \rangle < 1\}. \quad (2.2)$$

If $\alpha \in \mathbb{K}_\infty$ is of the form (2.1), and M is an integer, we put

$$[\alpha]_M = \sum_{i=M}^n a_i t^i \quad (2.3)$$

and $\{\alpha\}_M = \alpha - [\alpha]_M$. Evidently $[\alpha]_0 \in \mathbb{A}$ and $\{\alpha\}_0 \in \mathbb{T}$. It is also convenient to write $\langle\langle \alpha \rangle\rangle = \langle\{\alpha\}_0\rangle$ for any $\alpha \in \mathbb{K}_\infty$. For any integers m and M with m positive, we extend the operators $[\cdot]_M$, $\{\cdot\}_M$ and $\langle\langle \cdot \rangle\rangle$ to \mathbb{K}_∞^m , by

$$[\boldsymbol{\alpha}]_M = ([\alpha_1]_M, \dots, [\alpha_m]_M),$$

$$\{\boldsymbol{\alpha}\}_M = (\{\alpha_1\}_M, \dots, \{\alpha_m\}_M)$$

and

$$\langle\langle \boldsymbol{\alpha} \rangle\rangle = \max \{\langle\langle \alpha_j \rangle\rangle : j = 1, \dots, m\}. \quad (2.4)$$

As usual, we write $e(z)$ for $e^{2\pi iz}$ when z is real. We need a similar exponential function on the function field analogue \mathbb{K}_∞ of \mathbb{R} . For any element α of \mathbb{K}_∞ given by (2.1), define its residue $\text{res } \alpha$ by

$$\text{res } \alpha = \begin{cases} a_{-1}, & \text{when } n \geq -1, \\ 0, & \text{when } n < -1. \end{cases} \quad (2.5)$$

Let $\text{tr} : \mathbb{F}_q \rightarrow \mathbb{F}_p$ be the trace map of the field extension \mathbb{F}_q over \mathbb{F}_p . Define the map $e_q : \mathbb{F}_q \rightarrow \mathbb{C}$ by $e_q(a) = e(\text{tr}(a)/p)$. We then define the exponential function $E : \mathbb{K}_\infty \rightarrow \mathbb{C}$ by

$$E(\alpha) = e_q(\text{res } \alpha). \quad (2.6)$$

Then for any $\gamma \in \mathbb{K}_\infty$, the map from \mathbb{K}_∞ to \mathbb{C}^\times given by $\alpha \mapsto E(\gamma\alpha)$ is an additive character on the additive group \mathbb{K}_∞ .

Unless indicated otherwise, all summations other than those in the shape $\sum_{a=\mu}^\nu$, where $\mu, \nu \in \mathbb{Z} \cup \{\pm\infty\}$, are over polynomials in \mathbb{A} or tuples of elements in \mathbb{A} . The set \mathbb{T} defined in (2.2) is a compact additive subgroup of \mathbb{K}_∞ , and thus possesses a unique Haar measure $d\alpha$. The Haar measure on \mathbb{T} is normalised so that $\int_{\mathbb{T}} d\alpha = 1$. From Lemma 1(f) of [10], we also have the standard orthogonality property that

$$\int_{\mathbb{T}} E(\alpha x) d\alpha = \begin{cases} 1, & \text{for } x = 0, \\ 0, & \text{for } x \in \mathbb{A} \setminus \{0\}. \end{cases} \quad (2.7)$$

Hence the set \mathbb{T} defined in (2.2) can be regarded as the function field analogue of \mathbb{R}/\mathbb{Z} . Equation (3) of [10] provides the useful relation

$$\int_{\langle \alpha \rangle < \widehat{m}^{-1}} d\alpha = \widehat{m}^{-1}, \quad \text{for all } m \in \mathbb{N}. \quad (2.8)$$

We can extend the Haar measure on \mathbb{T} to a (unique) translation-invariant measure on \mathbb{K}_∞ by countable additivity. Indeed, let m be a positive integer. The set of all $\alpha \in \mathbb{K}_\infty$ with $\langle \alpha \rangle < \widehat{m}$ is the disjoint union of the translates $a_0 + a_1 t + \dots + a_{m-1} t^{m-1} + \mathbb{T}$ of \mathbb{T} , with a_0, \dots, a_{m-1} ranging over \mathbb{F}_q . Using this together with the normalisation $\int_{\mathbb{T}} d\alpha = 1$, we have

$$\int_{\langle \alpha \rangle < \widehat{m}} d\alpha = \sum_{\langle x \rangle < \widehat{m}} \int_{\mathbb{T}+x} d\alpha = \sum_{a_0, \dots, a_{m-1} \in \mathbb{F}_q} \int_{\mathbb{T}} d\alpha = \widehat{m}. \quad (2.9)$$

The measures on \mathbb{T} and \mathbb{K}_∞ also extend easily to product measures on the respective D -fold Cartesian products \mathbb{T}^D and \mathbb{K}_∞^D , for any positive integer D .

We examine the weak approximation property by introducing the quantities $\boldsymbol{\xi}$, \mathbf{b} , h and N . Let $\boldsymbol{\xi} \in \mathbb{T}^s$. Let h be a monic polynomial in \mathbb{A} . Let $\mathbf{b} \in \mathbb{A}^s$ with $\langle \mathbf{b} \rangle < \langle h \rangle$. Take N to be a non-negative integer that is bounded in terms of $\langle h \rangle$. Write W for the affine algebraic set given by the equations

$$\mathbf{F}(h\mathbf{x} + \mathbf{b}) = \mathbf{0}. \quad (2.10)$$

Moreover, let \mathcal{B} be the hypercube defined by

$$\mathcal{B} = \mathcal{B}_N = \left\{ \boldsymbol{\alpha} \in \mathbb{T}^s : \langle \boldsymbol{\alpha} - \boldsymbol{\xi} \rangle < \widehat{N}^{-1} \right\}. \quad (2.11)$$

We call \widehat{N}^{-1} the *sidelength* of the hypercube \mathcal{B} . Until §9, we regard h and \mathbf{b} as fixed and suppress mention of the dependence of \mathcal{B} on N .

Let n be a monic polynomial in \mathbb{A} with degree P . When m is a monic polynomial in \mathbb{A} and $\mathcal{E} \subseteq \mathbb{T}^s$, we write $\mathbf{x} \in m\mathcal{E}$ to mean that $m^{-1}\mathbf{x} \in \mathcal{E}$. In addition, when this notation appears in a summation, the \mathbf{x} involved are taken

to be integral points in $m\mathcal{E}$. When m is as above, $\alpha \in \mathbb{K}_\infty^R$ and $\mathcal{E} \subseteq \mathbb{T}^s$, define the generating function

$$T(\alpha; m, \mathcal{E}) = \sum_{\mathbf{x} \in m\mathcal{E}} E(\alpha \cdot \mathbf{F}(h\mathbf{x} + \mathbf{b})). \quad (2.12)$$

In particular, write $T(\alpha) = T(\alpha; n, \mathcal{B})$. For any measurable subset Z of \mathbb{K}_∞^R , write

$$\rho_{h,\mathbf{b}}(n; Z) = \int_Z T(\alpha) \, d\alpha. \quad (2.13)$$

It follows from (2.7) that the number $\rho_{h,\mathbf{b}}(n)$ of integral solutions \mathbf{x} of (2.10) with $\mathbf{x} \in n\mathcal{B}$ is given by

$$\rho_{h,\mathbf{b}}(n) = \rho_{h,\mathbf{b}}(n; \mathbb{T}^R). \quad (2.14)$$

Before our analysis of the quantity $\rho_{h,\mathbf{b}}(n)$, a few comments on our notation are in order. Recall the Vinogradov notations $O(\cdot)$, \ll , \gg and \asymp . With the exception of Theorem 1.1, the implicit constants in each application of these symbols depend at most on s , d , R , h , \mathbf{b} and the coefficients of the forms \mathbf{F} of interest. The symbol ϖ , with or without subscripts, always denotes a monic irreducible polynomial in \mathbb{A} . For any non-zero $x \in \mathbb{A}$, we write $\varpi^l \parallel x$ to mean that $\varpi^l \mid x$ but $\varpi^{l+1} \nmid x$. The summation notation \sum^\dagger is reserved for sums over *monic* polynomials in \mathbb{A} . For any field L , we denote its algebraic closure by \overline{L} . Given any polynomial $H \in L[x_1, \dots, x_s]$, the symbol H^* denotes the form given by the sum of the terms of highest degree in $H(x_1, \dots, x_s)$. For any R -tuple $\mathbf{H} = (H_1, \dots, H_R)$ of polynomials in $L[x_1, \dots, x_s]$, the symbol $X_{\mathbf{H}}$ denotes the affine algebraic set over \overline{L} given by the equations $H_i^*(x_1, \dots, x_s) = 0$, where $i = 1, \dots, R$. For any positive integer D , write I_D for the D -dimensional identity matrix. Finally, it is worthwhile to comment on the abundance of vector notations for vectors of various dimensions arising in this paper. The interpretation of the dimension of each vector involved depends purely on the context in which it lies.

3. WEYL'S INEQUALITY

The aim of this section is to derive a generic upper bound for exponential sums of the form given in (2.12). Throughout this section, we fix a subset $\mathcal{E} \subseteq \mathbb{T}^s$ and a monic polynomial $m \in \mathbb{A}$ with degree Q , which we think of as tending to infinity. We also abbreviate the exponential sum $T(\alpha; m, \mathcal{E})$ to $U(\alpha)$, and write $\mathbf{G}(\mathbf{x}) = \mathbf{F}(h\mathbf{x} + \mathbf{b})$. Then

$$U(\alpha) = \sum_{\mathbf{x} \in m\mathcal{E}} E(\alpha \cdot \mathbf{F}(h\mathbf{x} + \mathbf{b})) = \sum_{\mathbf{x} \in m\mathcal{E}} E(\alpha \cdot \mathbf{G}(\mathbf{x})). \quad (3.1)$$

The degree- d part G_i^* of each polynomial G_i can be expressed as

$$G_i^*(\mathbf{x}) = \sum_{j_1, \dots, j_d=1}^s a_{j_1, \dots, j_d}^{(i)} x_{j_1} \dots x_{j_d}, \quad (3.2)$$

where the coefficients $a_{j_1, \dots, j_d}^{(i)}$ are in \mathbb{A} and symmetric with respect to permutations on their subscripts, for each $i = 1, \dots, R$. Without loss of generality, we assume that for each such i , the coefficients $a_{j_1, \dots, j_d}^{(i)}$ are coprime. It is also convenient to introduce their associated multilinear forms

$$\Psi_j^{(i)}(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(d-1)}) = d! \sum_{j_1, \dots, j_{d-1}=1}^s a_{j_1, \dots, j_{d-1}, j}^{(i)} x_{j_1}^{(1)} \dots x_{j_{d-1}}^{(d-1)} \quad (3.3)$$

for $i = 1, \dots, R$ and $j = 1, \dots, s$. We denote by $\Psi = \Psi(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(d-1)})$ the matrix $(\Psi_j^{(i)})$, with $i = 1, \dots, R$ and $j = 1, \dots, s$.

The following lemma, which we do not prove, is the function field analogue of Lemma 2.1 of [2].

Lemma 3.1. *For every $\alpha \in \mathbb{K}_\infty^R$, we have*

$$|U(\alpha)|^{2^{d-1}} \leq \widehat{Q}^{(2^{d-1}-d)s} \sum_{\langle \mathbf{x}^{(1)} \rangle, \dots, \langle \mathbf{x}^{(d-1)} \rangle < \widehat{Q}} \prod_{j=1}^s |\Upsilon_j|,$$

where $\Upsilon_j = \Upsilon_j(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(d-1)}; \alpha)$ is given by

$$\Upsilon_j = \sum_{\langle x \rangle < \widehat{Q}} E \left(\sum_{i=1}^R \alpha_i \Psi_j^{(i)}(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(d-1)}) x \right). \quad (3.4)$$

Proof. The argument used in obtaining the required estimate is standard in circle method literature. We refer the reader to the proof of Lemma 11.1 of [15] for the underlying argument. \square

Note that when $p \leq d$, the factor $d!$ forces the multilinear forms $\Psi_j^{(i)}$ to vanish identically. Hence, the derivation of a non-trivial upper bound for $U(\alpha)$ inevitably imposes the condition that $p > d$.

Before proceeding further, we quote the following useful lemma.

Lemma 3.2. *Let $\gamma \in \mathbb{K}_\infty$, and let c be a non-negative integer. Then*

$$\sum_{\langle x \rangle < \widehat{c}} E(\gamma x) = \begin{cases} \widehat{c}, & \text{if } \langle\langle \gamma \rangle\rangle < \widehat{c}^{-1}, \\ 0, & \text{otherwise.} \end{cases}$$

Proof. This is Lemma 7 of [10]. \square

Let η be a parameter in the interval $(0, 1)$ to be specified. Recall the notation $\langle\langle \cdot \rangle\rangle$ given in (2.4). For any $v = 0, 1, \dots, d-1$, let $N_\eta^{(v)}(\alpha)$ denote the number of $(d-1)$ -tuples $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(d-1)}) \in (\mathbb{A}^s)^{d-1}$ with

$$\langle \mathbf{x}^{(1)} \rangle, \dots, \langle \mathbf{x}^{(v)} \rangle < \widehat{Q}^\eta \quad \text{and} \quad \langle \mathbf{x}^{(v+1)} \rangle, \dots, \langle \mathbf{x}^{(d-1)} \rangle < \widehat{Q}, \quad (3.5)$$

such that

$$\langle\langle \alpha \Psi(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(d-1)}) \rangle\rangle < \widehat{Q}^{-v-1+v\eta}. \quad (3.6)$$

The following corollary majorises $U(\alpha)$ in terms of the quantity $N_\eta^{(0)}(\alpha)$.

Corollary 3.3. *Whenever $p > d$, we have*

$$|U(\boldsymbol{\alpha})|^{2^{d-1}} \leq \widehat{Q}^{(2^{d-1}-d+1)s} N_{\eta}^{(0)}(\boldsymbol{\alpha}).$$

Proof. By the previous lemma, the sum defining Υ_j in (3.4) is equal to \widehat{Q} when

$$\left\langle\left\langle \sum_{i=1}^R \alpha_i \Psi_{j_d}^{(i)}(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(d-1)}) \right\rangle\right\rangle < \widehat{Q}^{-1},$$

and zero otherwise. Using the definition of $N_{\eta}^{(0)}(\boldsymbol{\alpha})$, we obtain the required result by putting this back into (3.4). \square

Our next goal is to provide the tools for bounding the quantity $N_{\eta}^{(0)}(\boldsymbol{\alpha})$ in terms of the number $N_{\eta}^{(d-1)}(\boldsymbol{\alpha})$. The ultimate aim of this manoeuvre is to choose η in such a manner as to further relate integral solutions of (3.6), with $v = d-1$, to singular integral points on $X_{\mathbf{G}}$. The tools in question are provided by results from the geometry of numbers. We refer the reader to Appendix B for the basic notions and properties of lattices defined over \mathbb{K}_{∞} .

In our later deliberations, we take Λ to be the $2s$ -dimensional lattice given by the equation $\mathbf{x} = \Lambda \mathbf{u}$, where \mathbf{u} runs over all the elements of \mathbb{A}^{2s} ,

$$\Lambda = \begin{pmatrix} t^{-b} I_s & 0 \\ t^b \gamma & t^b I_s \end{pmatrix}, \quad (3.7)$$

b is an integer and $\gamma = (\gamma_{i,j})_{1 \leq i,j \leq s}$ is a symmetric matrix with entries in \mathbb{K}_{∞} . Let M be the adjoint lattice of Λ , with points $\mathbf{y} = M \mathbf{v}$, where \mathbf{v} runs over all the points in \mathbb{A}^{2s} . Then the underlying matrix of M is

$$M = \begin{pmatrix} t^b I_s & -t^b \gamma \\ 0 & t^{-b} I_s \end{pmatrix}.$$

We can make M essentially the same as Λ by first changing the signs of v_{s+1}, \dots, v_{2s} , then changing those of y_{s+1}, \dots, y_{2s} , then interchanging the two sets of variables v_1, \dots, v_s and v_{s+1}, \dots, v_{2s} , and finally interchanging the two sets of variables y_1, \dots, y_s and y_{s+1}, \dots, y_{2s} . Hence the two lattices Λ and M can be regarded as the same lattice. Write $\widehat{R}_1, \dots, \widehat{R}_{2s}$ for the successive minima of Λ . Lemma B.6 thus implies that

$$\widehat{R}_{\nu} \widehat{R}_{2s-\nu+1} = 1 \quad \text{for } \nu = 1, \dots, s. \quad (3.8)$$

The following lemma concerns the number of integral solutions of a system of linear inequalities defined over \mathbb{K}_{∞} .

Lemma 3.4. *For $i = 1, \dots, s$, let $L_i(\mathbf{u})$ be the linear forms in u_1, \dots, u_s given by the equations*

$$L_i(\mathbf{u}) = \gamma_{i,1} u_1 + \dots + \gamma_{i,s} u_s, \quad (3.9)$$

where $\gamma_{i,j} \in \mathbb{K}_{\infty}$ with $\gamma_{i,j} = \gamma_{j,i}$ for all $i, j = 1, \dots, s$. When $a, Z \in \mathbb{R}$, let $M(a; Z)$ denote the number of $u_1, \dots, u_{2s} \in \mathbb{A}$ for which

$$\langle u_1 \rangle, \dots, \langle u_s \rangle < \widehat{a} \widehat{Z}$$

and

$$\langle L_i(\mathbf{u}) + u_{s+i} \rangle < \widehat{Z}/\widehat{a} \quad \text{for all } i = 1, \dots, s.$$

Then whenever $\widehat{Z}_1 \leq \widehat{Z}_2 \leq 1$, we have

$$\frac{M(a; Z_1)}{M(a; Z_2)} \gg \left(\frac{\widehat{Z}_1}{\widehat{Z}_2} \right)^s.$$

Proof. Take $\gamma = (\gamma_{i,j})_{1 \leq i,j \leq s}$, and let Λ be the lattice given by (3.7) with $b = [a]$. On comparing (3.9) with (3.7), and using the inequalities $b - 1 < a \leq b$, we can bound $M(a; Z)$ above and below by the inequalities

$$\# \left\{ \mathbf{x} \in \Lambda : \langle \mathbf{x} \rangle < q^{-1} \widehat{Z} \right\} \leq M(a; Z) \leq \# \left\{ \mathbf{x} \in \Lambda : \langle \mathbf{x} \rangle < q \widehat{Z} \right\}.$$

Put $R_0 = -\infty$, so that $\widehat{R}_0 = 0$. Since $\widehat{Z}_1 \leq \widehat{Z}_2$, we have

$$\widehat{R}_\mu \leq q^{-1} \widehat{Z}_1 < \widehat{R}_{\mu+1} \quad (3.10)$$

and

$$\widehat{R}_\nu \leq q \widehat{Z}_2 < \widehat{R}_{\nu+1},$$

for some $\mu, \nu \in \{0, \dots, 2s\}$ with $\mu \leq \nu$, as a result of our assumption that $\widehat{Z}_1 \leq \widehat{Z}_2$. By Lemma B.5, we have

$$M(a; Z_1) \geq \# \left\{ \mathbf{x} \in \Lambda : \langle \mathbf{x} \rangle < q^{-1} \widehat{Z}_1 \right\} = \prod_{\omega=1}^{\mu} (q^{-1} \widehat{Z}_1 / \widehat{R}_\omega) \quad (3.11)$$

and

$$M(a; Z_2) \leq \# \left\{ \mathbf{x} \in \Lambda : \langle \mathbf{x} \rangle < q \widehat{Z}_2 \right\} = \prod_{\omega=1}^{\nu} (q \widehat{Z}_2 / \widehat{R}_\omega). \quad (3.12)$$

Dividing (3.11) by (3.12) gives

$$\frac{M(a; Z_1)}{M(a; Z_2)} \geq q^{-\mu-\nu} \left(\prod_{\omega=\mu+1}^{\nu} \frac{\widehat{R}_\omega}{\widehat{Z}_1} \right) \left(\frac{\widehat{Z}_1}{\widehat{Z}_2} \right)^\nu \gg \left(\prod_{\omega=\mu+1}^{\nu} \frac{\widehat{R}_\omega}{\widehat{Z}_1} \right) \left(\frac{\widehat{Z}_1}{\widehat{Z}_2} \right)^\nu. \quad (3.13)$$

Suppose $\nu \leq s$. Using the second inequality in (3.10), the assumption $\widehat{Z}_1 \leq \widehat{Z}_2$ and the fact that $\widehat{R}_{\mu+1} \leq \dots \leq \widehat{R}_\nu$, we get

$$\frac{M(a; Z_1)}{M(a; Z_2)} \gg \prod_{\omega=\mu+1}^{\nu} \frac{\widehat{R}_\omega}{\widehat{R}_{\mu+1}} \left(\frac{\widehat{Z}_1}{\widehat{Z}_2} \right)^s \geq \left(\frac{\widehat{Z}_1}{\widehat{Z}_2} \right)^s.$$

We have thus proved the lemma in the case where $\nu \leq s$.

Suppose $\mu \leq s < \nu$. We can rewrite (3.13) in the form

$$\frac{M(a; Z_1)}{M(a; Z_2)} \gg \left(\frac{\widehat{Z}_1^\mu (\widehat{R}_{\mu+1} \dots \widehat{R}_s)}{\widehat{Z}_2^s} \right) \left(\frac{\widehat{R}_{s+1} \dots \widehat{R}_\nu}{\widehat{Z}_2^{\nu-s}} \right). \quad (3.14)$$

From (3.8), we have $\widehat{R}_s \widehat{R}_{s+1} = 1$. Since $\widehat{R}_s \leq \widehat{R}_{s+1}$, it follows that $\widehat{R}_s \leq 1 \leq \widehat{R}_{s+1}$. Since $\nu > s \geq \mu$, we also have

$$\widehat{R}_\nu \geq \dots \geq \widehat{R}_{s+1} \geq 1 \quad \text{and} \quad \widehat{R}_s \geq \dots \geq \widehat{R}_{\mu+1} \geq \widehat{Z}_1. \quad (3.15)$$

This, together with (3.14) and the assumption that $\widehat{Z}_2 \leq 1$, gives rise to

$$\frac{M(a; Z_1)}{M(a; Z_2)} \gg \frac{\widehat{Z}_1^\mu (\widehat{Z}_1^{s-\mu})}{\widehat{Z}_2^s} = \left(\frac{\widehat{Z}_1}{\widehat{Z}_2} \right)^s,$$

as required.

Finally let $s < \mu \leq \nu$. We first rewrite (3.13) as

$$\frac{M(a; Z_1)}{M(a; Z_2)} \gg \left(\frac{\widehat{Z}_1}{\widehat{Z}_2} \right)^s \frac{\widehat{Z}_1^{\mu-s} \widehat{R}_{\mu+1} \dots \widehat{R}_\nu}{\widehat{Z}_2^{\nu-s}}.$$

By the assumption that $\widehat{Z}_2 \leq 1$, this gives

$$\frac{M(a; Z_1)}{M(a; Z_2)} \gg \left(\frac{\widehat{Z}_1}{\widehat{Z}_2} \right)^s \widehat{Z}_1^{\mu-s} \widehat{R}_{\mu+1} \dots \widehat{R}_\nu. \quad (3.16)$$

Since now $\mu > s$, the left inequality in (3.10) further implies that

$$\widehat{Z}_1^{\mu-s} \geq \widehat{R}_\mu^{\mu-s} \geq \widehat{R}_\mu \dots \widehat{R}_{s+1}.$$

Also, since $\nu > s$, the first assertion in (3.15) still applies in the present case. Using this information in (3.16) thus gives

$$\frac{M(a; Z_1)}{M(a; Z_2)} \gg \left(\frac{\widehat{Z}_1}{\widehat{Z}_2} \right)^s \widehat{R}_{s+1} \dots \widehat{R}_\mu \widehat{R}_{\mu+1} \dots \widehat{R}_\nu \gg \left(\frac{\widehat{Z}_1}{\widehat{Z}_2} \right)^s,$$

as required. \square

Recall the quantity $N_\eta^{(v)}(\boldsymbol{\alpha})$ defined before Corollary 3.3. The following lemma re-expresses the upper bound in Corollary 3.3 by majorising the quantity $N_\eta^{(0)}(\boldsymbol{\alpha})$ in terms of $N_\eta^{(d-1)}(\boldsymbol{\alpha})$, via successive applications of Lemma 3.4.

Lemma 3.5. *For any $\boldsymbol{\alpha} \in \mathbb{K}_\infty^R$ and $\eta \in [0, 1]$, we have*

$$|U(\boldsymbol{\alpha})|^{2^{d-1}} \ll (\widehat{Q}^s)^{2^{d-1}-\eta(d-1)} N_\eta^{(d-1)}(\boldsymbol{\alpha}).$$

Proof. Courtesy of Corollary 3.3, the conclusion of this lemma holds if we establish for each $\boldsymbol{\alpha} \in \mathbb{K}_\infty^R$ that

$$N_\eta^{(d-1)}(\boldsymbol{\alpha}) \gg \widehat{Q}^{s(d-1)(\eta-1)} N_\eta^{(0)}(\boldsymbol{\alpha}).$$

We first show that for each $v = 1, \dots, d-1$, we have

$$N_\eta^{(v)}(\boldsymbol{\alpha}) \gg \widehat{Q}^{s(\eta-1)} N_\eta^{(v-1)}(\boldsymbol{\alpha}). \quad (3.17)$$

Fix such a v together with $\mathbf{x}^{(u)} \in \mathbb{A}^s$ ($1 \leq u \leq d-1, u \neq v$) which satisfy (3.5). We apply Lemma 3.4 with

$$L_j(\mathbf{x}^{(v)}) = \sum_{i=1}^R \alpha_i \Psi_j^{(i)}(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(v)}, \dots, \mathbf{x}^{(d-1)}) \quad \text{for each } j = 1, \dots, s.$$

This application is justified as long as these $L_j(\mathbf{x}^{(v)})$ satisfy the symmetry property as alluded to in that lemma. More precisely, we need to verify that

for every $j, j_v = 1, \dots, s$, the coefficient of $x_{j_v}^{(v)}$ in $L_j(\mathbf{x}^{(v)})$ is unchanged by interchanging j with j_v . Indeed, from (3.3), we see that

$$\begin{aligned} L_j(\mathbf{x}^{(v)}) &= d! \sum_{i=1}^R \alpha_i \sum_{j_1, \dots, j_{d-1}=1}^s a_{j_1, \dots, j_{d-1}, j}^{(i)} x_{j_1}^{(1)} \dots x_{j_{d-1}}^{(d-1)} \\ &= d! \sum_{i=1}^R \alpha_i \sum_{j_v=1}^s x_{j_v}^{(v)} \sum_{\substack{1 \leq j_u \leq s \\ (1 \leq u \leq d-1, \\ u \neq v)}} a_{j_1, \dots, j_{d-1}, j}^{(i)} x_{j_1}^{(1)} \dots x_{j_{v-1}}^{(v-1)} x_{j_{v+1}}^{(v+1)} \dots x_{j_{d-1}}^{(d-1)}. \end{aligned}$$

The coefficient of $x_{j_v}^{(v)}$ in the expansion for $L_j(\mathbf{x}^{(v)})$ is thus

$$d! \sum_{i=1}^R \alpha_i \sum_{\substack{1 \leq j_u \leq s \\ (1 \leq u \leq d-1, \\ u \neq v)}} a_{j_1, \dots, j_{d-1}, j}^{(i)} x_{j_1}^{(1)} \dots x_{j_{v-1}}^{(v-1)} x_{j_{v+1}}^{(v+1)} \dots x_{j_{d-1}}^{(d-1)}.$$

Since the coefficients $a_{j_1, \dots, j_{d-1}, j}^{(i)}$ are symmetric with respect to any permutation on their indices, it follows that the coefficient of $x_{j_v}^{(v)}$ in the expansion for $L_j(\mathbf{x}^{(v)})$ above is symmetric in the indices j and j_v , as desired. We can therefore estimate the number M_1 of $\mathbf{x}^{(v)} \in \mathbb{A}^s$ with $\langle \mathbf{x}^{(v)} \rangle < \widehat{Q}^\eta$ which satisfy (3.6), in terms of the number M_2 of $\mathbf{x}^{(v)} \in \mathbb{A}^s$ with $\langle \mathbf{x}^{(v)} \rangle < \widehat{Q}$ which also satisfy (3.6), but with v replaced by $v-1$ in the last quoted relation. In the notation of that lemma, we have $M_1 = M(a; Z_1)$ and $M_2 = M(a; Z_2)$, with

$$\widehat{a} = \widehat{Q}^{(v+1-(v-1)\eta)/2}, \quad \widehat{Z}_1 = \widehat{Q}^{(v+1)(\eta-1)/2} \quad \text{and} \quad \widehat{Z}_2 = \widehat{Q}^{(v-1)(\eta-1)/2}.$$

Since $\eta \leq 1$, we have $\widehat{Z}_1 \leq \widehat{Z}_2 \leq 1$, so Lemma 3.4 is applicable. An application of that lemma thus yields

$$M_1 \gg \left(\frac{\widehat{Q}^{(v+1)(\eta-1)/2}}{\widehat{Q}^{(v-1)(\eta-1)/2}} \right)^s M_2 = \widehat{Q}^{s(\eta-1)} M_2.$$

Now summing over all the $\mathbf{x}^{(u)}$ ($1 \leq u \leq d-1, u \neq v$) which satisfy (3.5), and recalling the definitions of M_1 and M_2 above, we obtain (3.17). Iterating on (3.17) thus leads to

$$N_\eta^{(d-1)}(\boldsymbol{\alpha}) \gg \widehat{Q}^{s(d-1)(\eta-1)} N_\eta^{(0)}(\boldsymbol{\alpha}),$$

as required. \square

Our remaining task is to relate the tuples counted by $N_\eta^{(d-1)}(\boldsymbol{\alpha})$ to the integral singular points on $X_{\mathbf{G}}$. Henceforth, the notation $\mathbf{a}/g \in \mathbb{K}^R$ means that g is a monic polynomial in \mathbb{A} , and $\mathbf{a} \in \mathbb{A}^R$ such that $(a_1, \dots, a_R, g) = 1$. We make use of the following lemma, which is a higher-dimensional generalisation of Lemma 2.3 of [8] to function fields. It provides conditions under which we can infer from a system of linear inequalities that the underlying matrix does not have full rank.

Lemma 3.6. *Let $M \in \mathbb{R}$, and $\alpha = \mathbf{a}/g + \beta$, where $\mathbf{a}/g \in \mathbb{K}^R$ and $\beta \in \mathbb{T}^R$. Suppose $\langle g\beta \rangle < \widehat{M}^{-1}$, and $\Phi = (\phi_{i,j})$ is an $R \times s$ -matrix with entries in \mathbb{A} such that $\langle \phi_{i,j} \rangle < \widehat{M}$ for all $i = 1, \dots, R$ and $j = 1, \dots, s$. Moreover, assume that $\langle\langle \alpha\Phi \rangle\rangle < \widehat{Y}^{-1}$, where $Y \in \mathbb{R}$ such that $\widehat{Y} > \langle g \rangle$. Then the rank of $\Phi \pmod{g}$ is less than R . In addition, if either*

$$\langle g \rangle \geq \widehat{M}^R \quad \text{or} \quad \langle g\alpha - \mathbf{a} \rangle \geq \widehat{Y}^{-1}\widehat{M}^{R-1},$$

then $\text{rank } \Phi < R$.

Proof. We first prove the first assertion of the lemma. The assumptions $\langle g\beta \rangle < \widehat{M}^{-1}$ and $\langle \phi_{i,j} \rangle < \widehat{M}$ ($1 \leq j \leq s, 1 \leq i \leq R$) imply that

$$\begin{aligned} \langle \beta\Phi \rangle &= \max_{1 \leq j \leq s} \left\langle \sum_{i=1}^R \beta_i \phi_{i,j} \right\rangle \leq \max_{1 \leq j \leq s} \max_{1 \leq i \leq R} \langle \beta_i \rangle \langle \phi_{i,j} \rangle \\ &< (\langle g \rangle^{-1} \widehat{M}^{-1}) \widehat{M} = \langle g \rangle^{-1}. \end{aligned} \tag{3.18}$$

Therefore $\beta\Phi \in \mathbb{T}^s$. We then deduce that

$$\{\mathbf{a}\Phi/g\}_0 = \{\alpha\Phi - \beta\Phi\}_0 = \{\alpha\Phi\}_0 - \beta\Phi.$$

Using the assumptions that $\langle\langle \alpha\Phi \rangle\rangle < \widehat{Y}^{-1}$ and $\langle g \rangle < \widehat{Y}$ in conjunction with (3.18), we get

$$\langle\langle \mathbf{a}\Phi/g \rangle\rangle \leq \max \{ \langle\langle \alpha\Phi \rangle\rangle, \langle \beta\Phi \rangle \} < \max \{ \widehat{Y}^{-1}, \langle g \rangle^{-1} \} = \langle g \rangle^{-1}.$$

This necessarily leads to $g \mid \mathbf{a}\Phi$. Since not all of the a_i are divisible by g , it follows from $g \mid \mathbf{a}\Phi$ that \mathbf{a} is a non-trivial solution to the homogeneous linear equation

$$\mathbf{a}\Phi \equiv \mathbf{0} \pmod{g}.$$

The matrix $\Phi \pmod{g}$ must therefore have less than the full rank R . This concludes the proof of the first assertion of the lemma.

We then establish the remaining assertion. Suppose $\text{rank } \Phi = R$. Then one of its $R \times R$ -determinants does not vanish. To ease the notation, we assume without loss of generality that this non-vanishing determinant is the leading minor of Φ . Let H denote the $R \times R$ -matrix formed from the top left corner of Φ . The first assertion of the lemma implies that every $R \times R$ -determinant of Φ , and in particular $\det H$, is divisible by g as a result. Also, every entry of Φ has size less than \widehat{M} , so $\langle \det H \rangle < \widehat{M}^R$. These two observations lead to

$$\langle g \rangle \leq \langle \det H \rangle < \widehat{M}^R.$$

Our work in the last paragraph shows that $g \mid \mathbf{a}\Phi$, and in particular g divides the first R coordinates of the s -dimensional vector $\mathbf{a}\Phi$. But these coordinates form the vector $\mathbf{a}H$. On recalling that $\alpha = \mathbf{a}/g + \beta$, we can write

$$\alpha H = \mathbf{A} + \Upsilon,$$

where $\mathbf{A} = \mathbf{a}H/g \in \mathbb{A}^R$ and $\Upsilon = \beta H \in \mathbb{K}_\infty^R$. Hence $g\alpha - \mathbf{a}$ satisfies the linear equation

$$(g\alpha - \mathbf{a})H = g\Upsilon.$$

By our assumption that $\det H \neq 0$, we can express $g\alpha - \mathbf{a}$ in terms of H , g and Υ by Cramér's rule. Since $g \mid \det H$ and each entry of H has size less than \widehat{M} , an application of Cramér's rule leads to the bound

$$\langle g\alpha - \mathbf{a} \rangle < \langle g/(\det H) \rangle \langle \Upsilon \rangle \widehat{M}^{R-1} \leq \langle \Upsilon \rangle \widehat{M}^{R-1}.$$

Finally, the earlier deductions that $\beta\Phi \in \mathbb{T}^s$ and $g|\mathbf{a}\Phi$, together with the assumption $\langle\langle \alpha\Phi \rangle\rangle < \widehat{Y}^{-1}$, implies that

$$\langle \Upsilon \rangle = \langle \beta H \rangle \leq \langle \beta\Phi \rangle = \langle\langle \beta\Phi \rangle\rangle = \langle\langle \alpha\Phi \rangle\rangle < \widehat{Y}^{-1}.$$

We have therefore shown that, if $\text{rank } \Phi = R$, then $\langle g \rangle < \widehat{M}^R$ and $\langle g\alpha - \mathbf{a} \rangle < \widehat{Y}^{-1} \widehat{M}^{R-1}$, contrary to the assumptions made in the latter half of the statement of the lemma. \square

Let $\alpha = \mathbf{a}/g + \beta$, where $\mathbf{a}/g \in \mathbb{K}^R$, and $\beta \in \mathbb{T}^R$. Our intermediate goal is to bound $N_\eta^{(d-1)}(\alpha)$ above by the number U_η of $(d-1)$ -tuples $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(d-1)}) \in \mathbb{A}^{s(d-1)}$ satisfying (3.5) with $v = d-1$, such that

$$\text{rank } \Psi(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(d-1)}) < R. \quad (3.19)$$

To this end, we choose η carefully so as to render Lemma 3.6 applicable. Write $\widehat{\kappa}$ for the maximum absolute value of the coefficients of the multilinear forms $\Psi_j^{(i)}$. We then apply that lemma with

$$\widehat{M} = \widehat{\kappa} \widehat{Q}^{(d-1)\eta}, \quad \widehat{Y} = \widehat{Q}^{d-(d-1)\eta},$$

and $\phi_{i,j} = \Psi_j^{(i)}(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(d-1)})$, where $(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(d-1)})$ varies over those $(d-1)$ -tuples that are counted by $N_\eta^{(d-1)}(\alpha)$. An examination of the hypotheses of Lemmata 3.5 and 3.6 reveals that they are both applicable if η satisfies $\eta \leq 1$,

$$\langle g\beta \rangle < \widehat{\kappa}^{-1} \widehat{Q}^{-(d-1)\eta}, \quad \langle g \rangle < \widehat{Q}^{d-(d-1)\eta},$$

and

$$\langle g \rangle \geq \min \left\{ \widehat{\kappa}^R \widehat{Q}^{R(d-1)\eta}, \widehat{\kappa}^{R-1} \widehat{Q}^{R(d-1)\eta} (\widehat{Q}^d \langle \beta \rangle)^{-1} \right\}.$$

The bound given by Lemma 3.5 is optimised by making η as large as possible. We thus choose η so that

$$\begin{aligned} & \widehat{Q}^{R(d-1)\eta} \\ &= q^{-1} \min \left\{ \widehat{Q}^{R(d-1)}, \widehat{\kappa}^{-R} \langle g\beta \rangle^{-R}, \widehat{Q}^{Rd} \langle g \rangle^{-R}, \langle g \rangle \widehat{\kappa}^{-R} \max \left\{ 1, \widehat{\kappa} \widehat{Q}^d \langle \beta \rangle \right\} \right\}. \end{aligned} \quad (3.20)$$

An application of Lemma 3.6, then yields

$$N_\eta^{(d-1)}(\alpha) \leq U_\eta. \quad (3.21)$$

Finally we relate the points counted by U_η to the singular points of the affine algebraic set $X_{\mathbf{G}}$. Let \mathcal{U} be the set in affine $s(d-1)$ -space given by the relation (3.19), and let \mathcal{D} be the diagonal in affine $s(d-1)$ -space given by the equations

$$\mathbf{x}^{(1)} = \dots = \mathbf{x}^{(d-1)}.$$

From (3.2) and (3.3), we deduce that when $i = 1, \dots, R$ and $j = 1, \dots, s$, we have

$$(d-1)! \partial_j G_i^*(\mathbf{x}) = d! \sum_{j_1, \dots, j_{d-1}=1}^s a_{j_1, \dots, j_{d-1}, j}^{(i)} x_{j_1} \dots x_{j_{d-1}} = \Psi_j^{(i)}(\mathbf{x}, \dots, \mathbf{x}).$$

So when $p > d$, we have $\mathcal{U} \cap \mathcal{D} = \text{Sing } X_{\mathbf{G}}$. Let σ temporarily denote the dimension of the singular locus of $X_{\mathbf{G}}$. By the affine dimension theorem (see Proposition I.7.1 of [6]), we get

$$\sigma = \dim(\mathcal{U} \cap \mathcal{D}) \geq \dim \mathcal{U} + \dim \mathcal{D} - s(d-1) = \dim \mathcal{U} - s(d-2).$$

Thus $\dim \mathcal{U} \leq \sigma + s(d-2)$. The argument which led to Lemma 3.1 of [2] implies that

$$U_\eta = \# \left\{ (\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(d-1)}) \in \mathcal{U} : \mathbf{x}^{(u)} \in \mathbb{A}^s, \langle \mathbf{x}^{(u)} \rangle < \widehat{Q}^\eta \text{ for } u = 1, \dots, d-1 \right\} \\ \ll (\widehat{Q}^\eta)^{\sigma+s(d-2)}.$$

This, together with (3.21) and Lemma 3.5, leads to the estimate

$$|U(\boldsymbol{\alpha})|^{2^{d-1}} \ll (\widehat{Q}^s)^{2^{d-1}-\eta(d-1)} U_\eta \\ \ll (\widehat{Q}^s)^{2^{d-1}-\eta(d-1)} (\widehat{Q}^\eta)^{\sigma+s(d-2)} \\ = \widehat{Q}^{2^{d-1}s} (\widehat{Q}^\eta)^{\sigma-s}.$$

Define the real number L by the equation

$$s - \sigma = 2^{d-1} L. \tag{3.22}$$

The above inequality can then be rewritten as

$$U(\boldsymbol{\alpha}) \ll \widehat{Q}^{s-L\eta} = \widehat{Q}^s (\widehat{Q}^{R(d-1)\eta})^{-L/(R(d-1))}.$$

The choice of η in (3.20) gives

$$U(\boldsymbol{\alpha}) \ll \widehat{Q}^s \min \left\{ \widehat{Q}^{R(d-1)}, \langle g\boldsymbol{\beta} \rangle^{-R}, \widehat{Q}^{Rd} \langle g \rangle^{-R}, \langle g \rangle (1 + \widehat{Q}^d \langle \boldsymbol{\beta} \rangle) \right\}^{-L/(R(d-1))}.$$

On recalling that $\boldsymbol{\beta} = \boldsymbol{\alpha} - \mathbf{a}/g$, we obtain the following lemma, commonly known as *Weyl's inequality* in circle method literature.

Lemma 3.7. *Let $\mathcal{E} \subseteq \mathbb{T}^s$. Set $\mathbf{G}(\mathbf{x}) = \mathbf{F}(h\mathbf{x} + \mathbf{b})$. Let L be given by (3.22). Let $\boldsymbol{\alpha} \in \mathbb{K}_\infty^R$, $\mathbf{a}/g \in \mathbb{K}^R$, and take m to be a monic polynomial in \mathbb{A} with large degree Q . Suppose $p > d$. Then*

$$U(\boldsymbol{\alpha}) \\ \ll \widehat{Q}^s \left(\frac{\widehat{Q}^R + \langle g \rangle^R + \widehat{Q}^{Rd} \langle g\boldsymbol{\alpha} - \mathbf{a} \rangle^R}{\widehat{Q}^{Rd}} + \frac{1}{\langle g \rangle + \widehat{Q}^d \langle g\boldsymbol{\alpha} - \mathbf{a} \rangle} \right)^{L/(R(d-1))}.$$

4. DISSECTION INTO MAJOR AND MINOR ARCS

The key step in counting solutions of the simultaneous equations (1.1) using the circle method is to dissect the torus \mathbb{T}^R into *major arcs*, and the complementary set, called the *minor arcs*. Informally speaking, the major arcs consist of those R -tuples in \mathbb{K}_∞^R that are simultaneously well approximated by R -tuples of rationals in \mathbb{K}^R whose common denominator has small degree. The remaining step in our analysis is to show that the contribution from the minor arcs to the underlying integral in (2.14) is negligible compared to that from the major arcs. This is shown in the remaining sections of this paper.

We need two slightly different types of major arcs in our work. When $\theta \in (0, 1)$, and $\mathbf{a}/g \in \mathbb{K}^R$ with $\langle \mathbf{a} \rangle < \langle g \rangle < \widehat{P}^{R(d-1)\theta}$, write

$$\mathfrak{M}(g, \mathbf{a}; \theta) = \left\{ \boldsymbol{\alpha} \in \mathbb{T}^R : \langle g\boldsymbol{\alpha} - \mathbf{a} \rangle < \widehat{\Delta}^{-1} \langle h \rangle^{-d} \widehat{P}^{-d+R(d-1)\theta} \right\}. \quad (4.1)$$

Let $\mathfrak{M}(\theta)$ denote the union of all these sets over the R -tuples of rationals \mathbf{a}/g described above. Also write $\mathfrak{m}(\theta)$ for the complement of $\mathfrak{M}(\theta)$ in \mathbb{T}^R . The sets $\mathfrak{M}(\theta)$, for various values of θ , are used as a tool for the pruning technique employed in obtaining a satisfactory minor arc estimate in the next section. Meanwhile, let θ_0 be a parameter in $(0, 1)$ to be specified in the following section. When $\mathbf{a}/g \in \mathbb{K}^R$ with $\langle \mathbf{a} \rangle < \langle g \rangle < \widehat{P}^{R(d-1)\theta_0}$, write

$$\mathfrak{N}(g, \mathbf{a}) = \left\{ \boldsymbol{\alpha} \in \mathbb{T}^R : \langle g\boldsymbol{\alpha} - \mathbf{a} \rangle < \widehat{\Delta}^{-1} \langle h \rangle^{-d} \langle g \rangle \widehat{P}^{-d+R(d-1)\theta_0} \right\}. \quad (4.2)$$

Let \mathfrak{N} denote the union of all the sets $\mathfrak{N}(g, \mathbf{a})$, over the \mathbf{a}/g described above. Also write $\mathfrak{n} = \mathbb{T}^R \setminus \mathfrak{N}$. The sets $\mathfrak{N}(g, \mathbf{a})$ are the major arcs employed in obtaining an asymptotic formula for the major arc contribution to the quantity in (2.14).

With this purpose in mind, we require a mild condition on θ_0 under which the sets $\mathfrak{N}(g, \mathbf{a})$ are disjoint. This is provided by the following lemma.

Lemma 4.1. *Let θ_0 be a positive number with $\theta_0 \leq d/(3R(d-1))$. Then whenever $\mathbf{a}^{(1)}/g_1, \mathbf{a}^{(2)}/g_2$ are distinct in \mathbb{K}^R , the sets $\mathfrak{N}(g_1, \mathbf{a}^{(1)})$ and $\mathfrak{N}(g_2, \mathbf{a}^{(2)})$ are disjoint.*

Proof. Suppose the conclusion of the lemma were false. Then we could find two distinct R -tuples $\mathbf{a}^{(1)}/g_1, \mathbf{a}^{(2)}/g_2 \in \mathbb{K}^R$, with $\langle \mathbf{a}^{(l)} \rangle < \langle g_l \rangle < \widehat{P}^{R(d-1)\theta_0}$ for $l = 1, 2$, such that there would exist $\boldsymbol{\alpha} \in \mathfrak{N}(g_1, \mathbf{a}^{(1)}) \cap \mathfrak{N}(g_2, \mathbf{a}^{(2)})$. The ultrametric property of the absolute value $\langle \cdot \rangle$ on \mathbb{K}_∞^R , together with (4.2), would then give

$$\begin{aligned} \widehat{P}^{-2R(d-1)\theta_0} &< \langle g_1 g_2 \rangle^{-1} \leq \langle \mathbf{a}^{(1)}/g_1 - \mathbf{a}^{(2)}/g_2 \rangle \\ &\leq \max \left\{ \langle \boldsymbol{\alpha} - \mathbf{a}^{(1)}/g_1 \rangle, \langle \boldsymbol{\alpha} - \mathbf{a}^{(2)}/g_2 \rangle \right\} \\ &< \widehat{P}^{-d+R(d-1)\theta_0}. \end{aligned}$$

This would lead to $1 < \widehat{P}^{-d+3R(d-1)\theta_0}$, which is a contradiction under our hypothesis on θ_0 . \square

We also need a crude upper bound on the measure of each of the sets $\mathfrak{M}(g, a; \theta)$. From (4.1) and (2.8), we deduce that

$$\begin{aligned}
\text{Vol } \mathfrak{M}(\theta) &\leq \sum_{\langle g \rangle < \widehat{P}^{R(d-1)\theta}}^{\dagger} \sum_{\substack{\langle \mathbf{a} \rangle < \langle g \rangle \\ (a_1, \dots, a_R, g)=1}} \text{Vol } \mathfrak{M}(g, \mathbf{a}; \theta) \\
&\leq \sum_{\langle g \rangle < \widehat{P}^{R(d-1)\theta}}^{\dagger} \langle g \rangle^R \prod_{i=1}^R \left(\int_{\langle \beta_i \rangle < \widehat{\Delta}^{-1} \langle h \rangle^{-d} \langle g \rangle^{-1} \widehat{P}^{-d+R(d-1)\theta}} d\beta_i \right) \\
&= \sum_{\langle g \rangle < \widehat{P}^{R(d-1)\theta}}^{\dagger} \langle g \rangle^R \prod_{i=1}^R (\widehat{\Delta}^{-1} \langle h \rangle^{-d} \langle g \rangle^{-1} \widehat{P}^{-d+R(d-1)\theta}) \\
&\leq \widehat{\Delta}^{-R} \langle h \rangle^{-Rd} \widehat{P}^{-Rd+R(R+1)(d-1)\theta}.
\end{aligned} \tag{4.3}$$

5. MINOR ARC ESTIMATE

This section is dedicated to bounding the minor arc contribution $\rho_{h,\mathbf{b}}(P; \mathbf{n})$ using the pruning argument employed by Birch in the proof of Lemma 4.4 of [2]. Let K be the real number defined by the relation

$$s - \dim \text{Sing } X = 2^{d-1} K. \tag{5.1}$$

We need the following function field analogue of the multidimensional Dirichlet approximation lemma.

Lemma 5.1. *Let R be a positive integer, $Y > 0$ and $\boldsymbol{\alpha} \in \mathbb{K}_{\infty}^R$. Then there exists an R -tuple $\mathbf{a}/g \in \mathbb{K}^R$ with $\langle g \rangle \leq \widehat{Y}$ such that $\langle g\boldsymbol{\alpha} - \mathbf{a} \rangle^R < \widehat{Y}^{-1}$.*

Proof. Take $N = \lfloor Y/R \rfloor$, so that $NR \leq Y$. The assertion of the lemma then follows once we find a polynomial $g \in \mathbb{A}$ in the shape

$$g = g_0 + g_1 t + \dots + g_{NR} t^{NR}, \tag{5.2}$$

such that

$$\langle\langle g\boldsymbol{\alpha}_i \rangle\rangle = \langle \{g\alpha_i\}_0 \rangle < \widehat{N}^{-1} \quad \text{for all } i = 1, \dots, R. \tag{5.3}$$

The last condition requires that the coefficients of t^{-l} ($1 \leq l \leq N$) in the expansions of $\{g\alpha_i\}_0$ vanish for all i . This gives rise to exactly NR equations in the $NR + 1$ variables g_0, \dots, g_{NR} . Further, a short computation reveals that these equations are actually homogeneous and linear. This guarantees the existence of a non-zero $(NR + 1)$ -tuple (g_0, \dots, g_{NR}) , with the property that the corresponding polynomial g in (5.2) satisfies (5.3). This completes the proof of this lemma. \square

We then derive an upper bound for the generating function $T(\boldsymbol{\alpha})$ using Lemma 3.7, in the case when $\boldsymbol{\alpha} \notin \mathfrak{M}(\theta)$, for various θ .

Lemma 5.2. *Let $\boldsymbol{\alpha} \in \mathbb{K}_{\infty}^R$ and $0 < \theta \leq d/((R+1)(d-1))$ such that $\boldsymbol{\alpha} \notin \mathfrak{M}(\theta)$. Then*

$$T(\boldsymbol{\alpha}) \ll \widehat{P}^{s-K\theta}.$$

Proof. Observe that for each $i = 1, \dots, R$, the degree- d part of the polynomial $F_i(h\mathbf{x} + \mathbf{b})$ in \mathbf{x} is $F_i(h\mathbf{x}) = h^d F_i(\mathbf{x})$. On recalling the equations defining X in (1.1), we see that the algebraic set generated by these degree- d parts coincide with X . On considering (3.22), we thus take $L = K$ in applying Lemma 3.7 to the generating function

$$T(\boldsymbol{\alpha}) = \sum_{\mathbf{x} \in n\mathcal{B}} E(\boldsymbol{\alpha} \cdot \mathbf{F}(h\mathbf{x} + \mathbf{b})).$$

Let $\boldsymbol{\alpha} \notin \mathfrak{M}(\theta)$, and Y be a positive parameter to be chosen. By Lemma 5.1, we can find an R -tuple $\mathbf{a}/g \in \mathbb{K}^R$ with $\langle g \rangle \leq \hat{Y}$, such that

$$\langle g\boldsymbol{\alpha} - \mathbf{a} \rangle^R < \hat{Y}^{-1}.$$

The assumption on $\boldsymbol{\alpha}$, together with the definition (4.1) of the sets $\mathfrak{M}(g, \mathbf{a}; \theta)$, implies that

$$\langle g \rangle + \hat{P}^d \langle g\boldsymbol{\alpha} - \mathbf{a} \rangle \gg \hat{P}^{R(d-1)\theta}. \quad (5.4)$$

Meanwhile, it follows from the information obtained from Lemma 5.1 that

$$\hat{P}^R + \langle g \rangle^R + \hat{P}^{Rd} \langle g\boldsymbol{\alpha} - \mathbf{a} \rangle^R \ll \hat{P}^R + \hat{Y}^R + \hat{P}^{Rd} \hat{Y}^{-1}.$$

To optimise this bound, we choose $\hat{Y} = \hat{P}^{Rd/(R+1)}$. This yields

$$\hat{P}^R + \langle g \rangle^R + \hat{P}^{Rd} \langle g\boldsymbol{\alpha} - \mathbf{a} \rangle^R \ll \hat{Y}^R = (\hat{P}^{Rd})^{R/(R+1)}.$$

Applying Lemma 3.7 with $L = K$, $m = n$ and $Q = P$, and using the last estimate along with (5.4) in the process, we obtain

$$\begin{aligned} T(\boldsymbol{\alpha}) &\ll \hat{P}^s \left(\frac{\hat{P}^R + \langle g \rangle^R + \hat{P}^{Rd} \langle g\boldsymbol{\alpha} - \mathbf{a} \rangle^R}{\hat{P}^{Rd}} + \frac{1}{\langle g \rangle + \hat{P}^d \langle g\boldsymbol{\alpha} - \mathbf{a} \rangle} \right)^{K/(R(d-1))} \\ &\ll \hat{P}^s \left((\hat{P}^{Rd})^{R/(R+1)-1} + \hat{P}^{-R(d-1)\theta} \right)^{K/R(d-1)} \\ &\ll \hat{P}^{s-K\theta}. \end{aligned}$$

Here the last line follows from our assumption on θ . This completes the proof of the lemma. \square

Evidently the bound given in this lemma is sharpest when $\theta = d/((R+1)(d-1))$. With this in mind, we devise our pruning argument for the minor arc contribution $\rho_{h,\mathbf{b}}(n; \mathbf{n})$ as follows. Let T be a sufficiently large integer, and $\theta_0, \dots, \theta_T$ be positive numbers such that

$$1/(2R(d-1)) = \theta_0 < \theta_1 < \dots < \theta_T = d/((R+1)(d-1)), \quad (5.5)$$

and

$$\theta_{r+1} - \theta_r < (1 - R(R+1)(d-1)/K)\theta_0/2 \quad \text{for all } r = 0, 1, \dots, T-1. \quad (5.6)$$

As elaborated in the next section, the choice for θ_0 in (5.5) is essentially the largest possible value for which several key results in our subsequent major arc analysis hold. In the meantime, choosing θ_0 to be as large as possible in this respect offers a superior estimate for $\rho_{h,\mathbf{b}}(n; \mathbf{n})$, as revealed shortly. The

condition (5.6) serves a similar purpose on the minor arcs \mathfrak{n} , and it is satisfied as long as T is large enough. The plan now is to decompose \mathfrak{n} as the disjoint union of sets

$$\mathfrak{n} = (\mathfrak{M}(\theta_0) \setminus \mathfrak{N}) \cup \left(\bigcup_{r=0}^{T-1} (\mathfrak{M}(\theta_{r+1}) \setminus \mathfrak{M}(\theta_r)) \right) \cup \mathfrak{m}(\theta_T).$$

This together with (2.13) gives

$$\begin{aligned} & \rho_{h,\mathbf{b}}(n; \mathfrak{n}) \\ &= \rho_{h,\mathbf{b}}(n; \mathfrak{M}(\theta_0) \setminus \mathfrak{N}) + \sum_{r=0}^{T-1} \rho_{h,\mathbf{b}}(n; \mathfrak{M}(\theta_{r+1}) \setminus \mathfrak{M}(\theta_r)) + \rho_{h,\mathbf{b}}(n; \mathfrak{m}(\theta_T)). \end{aligned} \quad (5.7)$$

The contribution from each term above can be bounded using (4.3) and Lemma 5.2 in conjunction. To ensure that the bound obtained on each segment is satisfactory, we assume that

$$K > R(R+1)(d-1), \quad (5.8)$$

and choose the θ_r ($0 \leq r \leq T-1$) such that

$$\delta_r = K\theta_r - R(R+1)(d-1)\theta_{r+1} > 0 \quad \text{for } r = 0, \dots, T-1. \quad (5.9)$$

Our choice of θ_T in (5.5), together with (5.8), implies that the quantity

$$\delta_T = K\theta_T - Rd \quad (5.10)$$

is positive. Finally write

$$\delta = \min_{0 \leq r \leq T} \delta_r. \quad (5.11)$$

The desired minor arc estimate is given in the following lemma.

Lemma 5.3. *Under the assumptions (5.8) and $p > d$, we have*

$$\rho_{h,\mathbf{b}}(n; \mathfrak{n}) \ll \widehat{P}^{s-Rd-\delta},$$

where

$$\delta > (K - R(R+1)(d-1))/(4R(d-1)).$$

Proof. We bound the contribution from each term in (5.7) in turn. From (4.1) and (4.2), we see that $\mathfrak{M}(g, \mathbf{a}; \theta_0) \subseteq \mathfrak{N}(g, \mathbf{a})$ for all $\mathbf{a}/g \in \mathbb{K}^R$ for which both sets of arcs are defined. Hence $\mathfrak{M}(\theta_0) \setminus \mathfrak{N}$ is empty, and $\rho_{h,\mathbf{b}}(n; \mathfrak{M}(\theta_0) \setminus \mathfrak{N}) = 0$ as a result. On recalling (2.13), and using (4.3), Lemma 5.2 and (5.9) in turn, we obtain

$$\begin{aligned} |\rho_{h,\mathbf{b}}(n; \mathfrak{M}(\theta_{r+1}) \setminus \mathfrak{M}(\theta_r))| &\leq \text{Vol } \mathfrak{M}(\theta_{r+1}) \sup_{\alpha \notin \mathfrak{M}(\theta_r)} |T(\alpha)| \\ &\ll \widehat{P}^{-Rd+R(R+1)(d-1)\theta_{r+1}} \widehat{P}^{s-K\theta_r} \\ &= \widehat{P}^{s-Rd-\delta_r}, \end{aligned}$$

when $0 \leq r \leq T-1$. Finally, on applying (2.13), Lemma 5.2 and (5.10) together, we get

$$|\rho_{h,\mathbf{b}}(n; \mathfrak{m}(\theta_T))| \leq \sup_{\alpha \in \mathfrak{m}(\theta_T)} |T(\alpha)| \ll \widehat{P}^{s-K\theta_T} = \widehat{P}^{s-Rd-\delta_T}.$$

The desired estimate for $\rho_{h,\mathbf{b}}(n; \mathfrak{n})$ thus follows by substituting these three estimates back into (5.7), and recalling (5.11).

It remains to establish the lower bound for δ asserted by the lemma. Using (5.10) together with the choice of θ_T in (5.5), we can rewrite δ_T as

$$\delta_T = Kd/((R+1)(d-1)) - Rd = (K - R(R+1)(d-1))\theta_T.$$

Furthermore, it follows from (5.9), (5.5) and (5.6) that when $0 \leq r \leq T-1$, we get

$$\begin{aligned} \delta_r &= -K(\theta_{r+1} - \theta_r) + (K - R(R+1)(d-1))\theta_{r+1} \\ &> -K(1 - R(R+1)(d-1)/K)\theta_0/2 + (K - R(R+1)(d-1))\theta_0 \\ &= (K - R(R+1)(d-1))\theta_0/2. \end{aligned}$$

On recalling the definition (5.11) of δ as well as the choice of θ_0 in (5.5), we therefore arrive at the lower bound

$$\delta > (K - R(R+1)(d-1))\theta_0/2 = (K - R(R+1)(d-1))/(4R(d-1)),$$

as required. \square

6. MAJOR ARC ANALYSIS

Now we turn our attention towards obtaining an asymptotic formula for the major arc contribution $\rho_{h,\mathbf{b}}(n; \mathfrak{N})$. The standard approach here is to express the generating function $T(\boldsymbol{\alpha})$ in terms of analytically well-behaved functions, making explicit use of the assumption that $\boldsymbol{\alpha} \in \mathfrak{N}$ in the process. Before we proceed, we introduce the necessary notations for the aforementioned functions. When $g \in \mathbb{A}$ and $\mathbf{a} \in \mathbb{A}^R$, write

$$S(g, \mathbf{a}) = \sum_{\langle \mathbf{x} \rangle < \langle g \rangle} E(\mathbf{a} \cdot \mathbf{F}(h\mathbf{x} + \mathbf{b})/g). \quad (6.1)$$

When $\boldsymbol{\gamma} \in \mathbb{K}_{\infty}^R$ and $\mathcal{E} \subseteq \mathbb{T}^s$, write

$$I(\boldsymbol{\gamma}; \mathcal{E}) = \int_{\mathcal{E}} E(\boldsymbol{\gamma} \cdot \mathbf{F}(\boldsymbol{\sigma})) \, d\boldsymbol{\sigma}. \quad (6.2)$$

Also write $\widehat{\Delta}$ for the maximum absolute value of the coefficients of the forms F_i ($i = 1, \dots, R$).

The following lemma gives the desired decomposition of the generating function $U(\boldsymbol{\alpha})$, given in (3.1), under rather generic assumptions on $\boldsymbol{\alpha}$.

Lemma 6.1. *Let $\boldsymbol{\alpha} \in \mathbb{K}_{\infty}^R$. Take \mathcal{E} to be a hypercube in \mathbb{T}^s with sidelength \widehat{M}^{-1} , for some non-negative integer M . Let m be a monic polynomial in \mathbb{A} with degree Q . Suppose $\mathbf{a}/g \in \mathbb{K}^R$ such that $\langle g \rangle < \widehat{Q}\widehat{M}^{-1}$ and*

$$\langle g\boldsymbol{\alpha} - \mathbf{a} \rangle < \widehat{\Delta}^{-1} \langle h \rangle^{-d} \widehat{Q}^{1-d}. \quad (6.3)$$

Then with $\boldsymbol{\beta} = \boldsymbol{\alpha} - \mathbf{a}/g$, we have

$$U(\boldsymbol{\alpha}) = \widehat{Q}^s \langle g \rangle^{-s} S(g, \mathbf{a}) I(h^d m^d \boldsymbol{\beta}; \mathcal{E}).$$

Proof. To express $U(\boldsymbol{\alpha})$ as asserted in the statement of the lemma, we first examine the summation condition $\mathbf{x} \in m\mathcal{E}$ on integral s -tuples \mathbf{x} . We express the hypercube \mathcal{E} as

$$\mathcal{E} = \left\{ \boldsymbol{\gamma} \in \mathbb{T}^s : \langle \boldsymbol{\gamma} - \boldsymbol{\zeta} \rangle < \widehat{M}^{-1} \right\}, \quad (6.4)$$

for some $\boldsymbol{\zeta} \in \mathbb{T}^s$. The summation condition on \mathbf{x} is then equivalent to

$$\langle \mathbf{x} - m\boldsymbol{\zeta} \rangle < \widehat{Q}\widehat{M}^{-1}.$$

In terms of the truncation operation defined in (2.3), this can be rewritten further as

$$\lfloor \mathbf{x} \rfloor_{Q-M} = \boldsymbol{\zeta}^*,$$

where

$$\boldsymbol{\zeta}^* = \lfloor m\boldsymbol{\zeta} \rfloor_{Q-M}. \quad (6.5)$$

Hence $\mathbf{x} \in \mathbb{A}^s \cap m\mathcal{E}$ if and only if $\mathbf{x} = \mathbf{w} + \boldsymbol{\zeta}^*$, for some unique $\mathbf{w} \in \mathbb{A}^s$ with $\langle \mathbf{w} \rangle < \widehat{Q}\widehat{M}^{-1}$. Since \mathbb{K}_∞^s is equipped with the non-archimedean absolute value given in (1.3), every point in \mathcal{E} is a centre of \mathcal{E} . Without loss of generality, we can choose our centre $\boldsymbol{\zeta}$ of \mathcal{E} in such a way that no coordinate of $\boldsymbol{\zeta}^*$ equals zero. We can then express $U(\boldsymbol{\alpha})$ as

$$U(\boldsymbol{\alpha}) = \sum_{\langle \mathbf{w} \rangle < \widehat{Q}\widehat{M}^{-1}} E(\boldsymbol{\alpha} \cdot \mathbf{F}(h(\mathbf{w} + \boldsymbol{\zeta}^*) + \mathbf{b})) \quad (6.6)$$

$$= \sum_{\langle \mathbf{w} \rangle < \widehat{Q}\widehat{M}^{-1}} E(\boldsymbol{\alpha} \cdot \mathbf{F}(h\mathbf{w} + \tilde{\boldsymbol{\zeta}})), \quad (6.7)$$

where

$$\tilde{\boldsymbol{\zeta}} = h\boldsymbol{\zeta}^* + \mathbf{b}. \quad (6.8)$$

Since $\widehat{Q}\widehat{M}^{-1} > \langle g \rangle$ by assumption, we can express each \mathbf{w} occurring in the sum (6.7) uniquely in the shape $\mathbf{w} = g\mathbf{y} + \mathbf{z}$, where $\mathbf{y}, \mathbf{z} \in \mathbb{A}^s$ with $\langle \mathbf{y} \rangle < \widehat{Q}\widehat{M}^{-1}\langle g \rangle^{-1}$ and $\langle \mathbf{z} \rangle < \langle g \rangle$. It follows from (6.7) that

$$\begin{aligned} U(\boldsymbol{\alpha}) &= \sum_{\langle \mathbf{z} \rangle < \langle g \rangle} \sum_{\langle \mathbf{y} \rangle < \widehat{Q}\widehat{M}^{-1}\langle g \rangle^{-1}} E(\boldsymbol{\beta} \cdot \mathbf{F}(h(g\mathbf{y} + \mathbf{z}) + \tilde{\boldsymbol{\zeta}})) E(\boldsymbol{\alpha} \cdot \mathbf{F}(h(g\mathbf{y} + \mathbf{z}) + \tilde{\boldsymbol{\zeta}})/g) \\ &= \sum_{\langle \mathbf{z} \rangle < \langle g \rangle} E(\boldsymbol{\alpha} \cdot \mathbf{F}(h\mathbf{z} + \tilde{\boldsymbol{\zeta}})/g) \sum_{\langle \mathbf{y} \rangle < \widehat{Q}\widehat{M}^{-1}\langle g \rangle^{-1}} E(\boldsymbol{\beta} \cdot \mathbf{F}(h(g\mathbf{y} + \mathbf{z}) + \tilde{\boldsymbol{\zeta}})). \end{aligned} \quad (6.9)$$

Here we have used the simple observation that $\mathbf{F}(h(g\mathbf{y} + \mathbf{z}) + \tilde{\boldsymbol{\zeta}}) \equiv \mathbf{F}(h\mathbf{z} + \tilde{\boldsymbol{\zeta}}) \pmod{g}$.

We aim to remove the \mathbf{z} -dependence in the inner sum in (6.9). To this end, we seek to establish that

$$E(\boldsymbol{\beta} \cdot \mathbf{F}(h(g\mathbf{y} + \mathbf{z}) + \tilde{\boldsymbol{\zeta}})) = E(\boldsymbol{\beta} \cdot \mathbf{F}(h\mathbf{y} + \tilde{\boldsymbol{\zeta}}))$$

whenever $\mathbf{z}, \mathbf{y} \in \mathbb{A}^s$ with $\langle \mathbf{z} \rangle < \langle g \rangle$ and $\langle \mathbf{y} \rangle < \widehat{Q}\widehat{M}^{-1}\langle g \rangle^{-1}$. From (2.5) and (2.6), it suffices to show that

$$\text{ord}(\beta \cdot (\mathbf{F}(h(g\mathbf{y} + \mathbf{z}) + \widetilde{\zeta}) - \mathbf{F}(h g \mathbf{y} + \widetilde{\zeta}))) < -1 \quad (6.10)$$

whenever $\langle \mathbf{y} \rangle < \widehat{Q}\widehat{M}^{-1}\langle g \rangle^{-1}$ and $\langle \mathbf{z} \rangle < \langle g \rangle$. Since each F_i is a form of degree d with coefficients having degree at most Δ , it follows that

$$\begin{aligned} & \text{ord}(F_i(h g \mathbf{y} + h \mathbf{z} + \widetilde{\zeta}) - F_i(h g \mathbf{y} + \widetilde{\zeta})) \\ & \leq \Delta + \max_{1 \leq u \leq s} \text{ord}(h z_u) + (d-1) \max_{1 \leq u \leq s} \text{ord}(h g y_u + \widetilde{\zeta}_u). \end{aligned} \quad (6.11)$$

On recalling (6.8), we see that

$$h g y_u + \widetilde{\zeta}_u = h(g y_u + \zeta_u^*) + b_u.$$

Further, since $\langle g y_u \rangle < \widehat{Q}\widehat{M}^{-1} \leq \langle \zeta_u^* \rangle$ for all $u = 1, \dots, s$, we see that $g y_u + \zeta_u^* \neq 0$ for each u . This together with $\langle \mathbf{b} \rangle < \langle h \rangle$ gives

$$\text{ord}(h g y_u + \widetilde{\zeta}_u) \leq \text{ord}(h(g y_u + \zeta_u^*)).$$

Using the assumption $\langle \mathbf{y} \rangle < \widehat{Q}\widehat{M}^{-1}\langle g \rangle^{-1}$, we obtain $\langle g \mathbf{y} + \zeta^* \rangle < \widehat{Q}$. It follows that

$$\text{ord}(h g y_u + \widetilde{\zeta}_u) \leq \text{ord } h + (Q-1), \quad (6.12)$$

for all $u = 1, \dots, s$. Also, since $\langle \mathbf{z} \rangle < \langle g \rangle$, we have

$$\text{ord}(h z_u) \leq \text{ord } h + (\text{ord } g - 1), \quad (6.13)$$

for any such u . Putting (6.12) and (6.13) into (6.11) thus gives

$$\begin{aligned} & \text{ord}(F_i(h g \mathbf{y} + h \mathbf{z} + \widetilde{\zeta}) - F_i(h g \mathbf{y} + \widetilde{\zeta})) \\ & \leq \Delta + \text{ord } h + (\text{ord } g - 1) + (d-1)(\text{ord } h + Q - 1) \\ & = \Delta + d \text{ord } h + (d-1)Q + \text{ord } g - d. \end{aligned}$$

The assumption (6.3) also implies that for any such i , we have

$$\text{ord } \beta_i \leq -\Delta - d \text{ord } h - \text{ord } g - (d-1)Q - 1 \quad \text{for all } i = 1, \dots, R.$$

The last two inequalities combined give rise to the upper bound

$$\text{ord}(\beta \cdot (\mathbf{F}(h(g\mathbf{y} + \mathbf{z}) + \widetilde{\zeta}) - \mathbf{F}(h g \mathbf{y} + \widetilde{\zeta}))) \leq -d - 1 < -1.$$

The desired relation (6.10) thus follows.

Having settled the claim (6.10), it follows from (6.9) that

$$\begin{aligned} & U(\alpha) \\ & = \left(\sum_{\langle \mathbf{z} \rangle < \langle g \rangle} E(\mathbf{a} \cdot \mathbf{F}(h \mathbf{z} + \widetilde{\zeta})/g) \right) \left(\sum_{\langle \mathbf{y} \rangle < \widehat{Q}\widehat{M}^{-1}\langle g \rangle^{-1}} E(\beta \cdot \mathbf{F}(h g \mathbf{y} + \widetilde{\zeta})) \right). \end{aligned} \quad (6.14)$$

Recalling (6.8), we can write $h \mathbf{z} + \widetilde{\zeta} = h \mathbf{z} + h \zeta^* + \mathbf{b}$. From (6.5) and our choice of ζ , we see that $\langle \zeta^* \rangle \geq \widehat{Q}\widehat{M}^{-1} > \langle g \rangle$. Hence we can write $\zeta^* = g \boldsymbol{\eta} + \boldsymbol{\phi}$, where $\boldsymbol{\eta}, \boldsymbol{\phi} \in \mathbb{A}^s$ with $\langle \boldsymbol{\phi} \rangle < \langle g \rangle$. Thus

$$h \mathbf{z} + \widetilde{\zeta} = h \mathbf{z} + h(g \boldsymbol{\eta} + \boldsymbol{\phi}) + \mathbf{b} \equiv h(\mathbf{z} + \boldsymbol{\phi}) + \mathbf{b} \pmod{g}.$$

As \mathbf{z} runs through all residue classes (mod g), so does $\mathbf{z} + \boldsymbol{\phi}$. On applying (6.1), the first sum in (6.14) can therefore be simplified as

$$\sum_{\langle \mathbf{z} \rangle < \langle g \rangle} E(\mathbf{a} \cdot \mathbf{F}(h\mathbf{z} + \tilde{\boldsymbol{\zeta}})/g) = \sum_{\langle \mathbf{z} \rangle < \langle g \rangle} E(\mathbf{a} \cdot \mathbf{F}(h\mathbf{z} + \mathbf{b})/g) = S(g, \mathbf{a}). \quad (6.15)$$

To simplify the second sum in (6.14), we average over all $\mathbf{z} \in \mathbb{A}^s$ with $\langle \mathbf{z} \rangle < \langle g \rangle$, use the claim before (6.10) again, put back $\mathbf{w} = g\mathbf{y} + \mathbf{z}$, and use (6.8). This gives

$$\begin{aligned} & \sum_{\langle \mathbf{y} \rangle < \widehat{Q}\widehat{M}^{-1}\langle g \rangle^{-1}} E(\boldsymbol{\beta} \cdot \mathbf{F}(h g \mathbf{y} + \tilde{\boldsymbol{\zeta}})) \\ &= \langle g \rangle^{-s} \sum_{\langle \mathbf{z} \rangle < \langle g \rangle} \sum_{\langle \mathbf{y} \rangle < \widehat{Q}\widehat{M}^{-1}\langle g \rangle^{-1}} E(\boldsymbol{\beta} \cdot \mathbf{F}(h(g\mathbf{y} + \mathbf{z}) + \tilde{\boldsymbol{\zeta}})) \\ &= \langle g \rangle^{-s} \sum_{\langle \mathbf{w} \rangle < \widehat{Q}\widehat{M}^{-1}} E(\boldsymbol{\beta} \cdot \mathbf{F}(h\mathbf{w} + \tilde{\boldsymbol{\zeta}})) \\ &= \langle g \rangle^{-s} \sum_{\langle \mathbf{w} \rangle < \widehat{Q}\widehat{M}^{-1}} E(\boldsymbol{\beta} \cdot \mathbf{F}(h(\mathbf{w} + \boldsymbol{\zeta}^*) + \mathbf{b})) \\ &= \langle g \rangle^{-s} U(\boldsymbol{\beta}). \end{aligned} \quad (6.16)$$

Here the last two lines follows from (6.7) and (6.6) respectively, with $\boldsymbol{\alpha} = \boldsymbol{\beta}$ in the cited equations.

To complete the proof of this lemma, it therefore remains to show that

$$U(\boldsymbol{\beta}) = \widehat{Q}^s I(h^d m^d \boldsymbol{\beta}; \mathcal{E}). \quad (6.17)$$

To achieve this, we use a similar argument to the one employed in establishing (6.10). Our first step is to show that for any $\boldsymbol{\sigma} \in \mathbb{T}^s$ and any integral s -tuple \mathbf{w} with $\langle \mathbf{w} \rangle < \widehat{Q}\widehat{M}^{-1}$, we have

$$\langle \boldsymbol{\beta} \cdot (\mathbf{F}(h(\mathbf{w} + \boldsymbol{\zeta}^* + \boldsymbol{\sigma}) + \mathbf{b}) - \mathbf{F}(h(\mathbf{w} + \boldsymbol{\zeta}^*) + \mathbf{b})) \rangle < q^{-1}. \quad (6.18)$$

This suffices to give the relation

$$E(\boldsymbol{\beta} \cdot \mathbf{F}(h(\mathbf{w} + \boldsymbol{\zeta}^* + \boldsymbol{\sigma}) + \mathbf{b})) = E(\boldsymbol{\beta} \cdot \mathbf{F}(h(\mathbf{w} + \boldsymbol{\zeta}^*) + \mathbf{b}))$$

for all such $\boldsymbol{\sigma}$ and \mathbf{w} . To prove (6.18), recall that each F_i is a form of degree d with coefficients having absolute value at most $\widehat{\Delta}$. This implies that

$$\langle F_i(h(\mathbf{w} + \boldsymbol{\zeta}^* + \boldsymbol{\sigma}) + \mathbf{b}) - F_i(h(\mathbf{w} + \boldsymbol{\zeta}^*) + \mathbf{b}) \rangle \leq \widehat{\Delta} \langle h\boldsymbol{\sigma} \rangle \langle h(\mathbf{w} + \boldsymbol{\zeta}^*) + \mathbf{b} \rangle^{d-1}.$$

On recalling (6.5) and that $\langle \mathbf{b} \rangle < \langle h \rangle$ and $0 < \langle \mathbf{w} + \boldsymbol{\zeta}^* \rangle < \widehat{Q}$, we see that for any \mathbf{w} and $\boldsymbol{\sigma}$ involved in (6.18), we have the estimate

$$\begin{aligned} \langle F_i(h(\mathbf{w} + \boldsymbol{\zeta}^* + \boldsymbol{\sigma}) + \mathbf{b}) - F_i(h(\mathbf{w} + \boldsymbol{\zeta}^*) + \mathbf{b}) \rangle &\leq \widehat{\Delta} \langle h\boldsymbol{\sigma} \rangle \langle h(\mathbf{w} + \boldsymbol{\zeta}^*) \rangle^{d-1} \\ &\leq \widehat{\Delta} (\langle h \rangle q^{-1}) (\langle h \rangle q^{-1} \widehat{Q})^{d-1} \\ &= \widehat{\Delta} \langle h \rangle^d q^{-d} \widehat{Q}^{d-1}. \end{aligned}$$

The condition (6.3) also gives $\langle \beta_i \rangle < \widehat{\Delta}^{-1} \langle h \rangle^{-d} \widehat{Q}^{1-d}$ for each $i = 1, \dots, R$. On combining these two degree estimates, we get

$$\begin{aligned} & \langle \beta_i (F_i(h(\mathbf{w} + \boldsymbol{\zeta}^* + \boldsymbol{\sigma}) + \mathbf{b}) - F_i(h(\mathbf{w} + \boldsymbol{\zeta}^*) + \mathbf{b})) \rangle \\ & \leq (\widehat{\Delta}^{-1} \langle h \rangle^{-d} \widehat{Q}^{1-d} q^{-1}) (\widehat{\Delta} \langle h \rangle^d q^{-d} \widehat{Q}^{d-1}) \\ & = q^{-d-1} < q^{-1}. \end{aligned}$$

This yields (6.18). Integrating both sides of (6.6) with respect to all $\boldsymbol{\sigma} \in \mathbb{T}^s$, and using the normalisation on \mathbb{T} and the relation displayed after (6.18), we can thus express $U(\boldsymbol{\beta})$ as

$$U(\boldsymbol{\beta}) = \int_{\mathbb{T}^s} \sum_{\langle \mathbf{w} \rangle < \widehat{Q} \widehat{M}^{-1}} E(\boldsymbol{\beta} \cdot \mathbf{F}(h(\mathbf{w} + \boldsymbol{\sigma} + \boldsymbol{\zeta}^*) + \mathbf{b})) \, d\boldsymbol{\sigma}.$$

As \mathbf{w} runs through all integral s -tuples with $\langle \mathbf{w} \rangle < \widehat{Q} \widehat{M}^{-1}$, and $\boldsymbol{\sigma}$ runs through all points in \mathbb{T}^s , the s -tuple $\boldsymbol{\omega} = \mathbf{w} + \boldsymbol{\sigma}$ runs through all real points with $\langle \boldsymbol{\omega} \rangle < \widehat{Q} \widehat{M}^{-1}$. This leads to

$$U(\boldsymbol{\beta}) = \int_{\langle \boldsymbol{\omega} \rangle < \widehat{Q} \widehat{M}^{-1}} E(\boldsymbol{\beta} \cdot \mathbf{F}(h(\boldsymbol{\omega} + \boldsymbol{\zeta}^*) + \mathbf{b})) \, d\boldsymbol{\omega}. \quad (6.19)$$

We plan to make the change of variables $\boldsymbol{\omega} = m\boldsymbol{\sigma}$ in the integral above. To accomplish this using Theorem A.1, we need to verify that the function

$$\boldsymbol{\omega} \mapsto E(\boldsymbol{\beta} \cdot \mathbf{F}(h(\boldsymbol{\omega} + \boldsymbol{\zeta}^*) + \mathbf{b}))$$

is continuous when $\langle \boldsymbol{\omega} \rangle < \widehat{Q} \widehat{M}^{-1}$. Indeed, note that this function takes only a discrete set of values in \mathbb{C} . Hence, to check the required continuity condition, it suffices to show that

$$E(\boldsymbol{\beta} \cdot \mathbf{F}(h(\boldsymbol{\omega} + \boldsymbol{\zeta}^*) + \mathbf{b})) = E(\boldsymbol{\beta} \cdot \mathbf{F}(h(\boldsymbol{\omega}' + \boldsymbol{\zeta}^*) + \mathbf{b}))$$

when $\boldsymbol{\omega}, \boldsymbol{\omega}' \in \mathbb{K}_\infty^s$ such that $\langle \boldsymbol{\omega} - \boldsymbol{\omega}' \rangle$ is sufficiently small. On recalling (2.6) and (2.5), we can reduce this task further to showing that

$$\langle \boldsymbol{\beta} \cdot (\mathbf{F}(h(\boldsymbol{\omega} + \boldsymbol{\zeta}^*) + \mathbf{b}) - \mathbf{F}(h(\boldsymbol{\omega}' + \boldsymbol{\zeta}^*) + \mathbf{b})) \rangle < q^{-1}, \quad (6.20)$$

when $\langle \boldsymbol{\omega} \rangle, \langle \boldsymbol{\omega}' \rangle < \widehat{Q} \widehat{M}^{-1}$ and $\langle \boldsymbol{\omega} - \boldsymbol{\omega}' \rangle$ is small enough. Observe that

$$(h(\boldsymbol{\omega} + \boldsymbol{\zeta}^*) + \mathbf{b}) - (h(\boldsymbol{\omega}' + \boldsymbol{\zeta}^*) + \mathbf{b}) = h(\boldsymbol{\omega} - \boldsymbol{\omega}').$$

Since each form F_i has degree d with coefficients having absolute value at most $\widehat{\Delta}$, it follows from this observation together with the inequalities $\langle \mathbf{b} \rangle < \langle h \rangle$ and $\langle \boldsymbol{\omega}' + \boldsymbol{\zeta}^* \rangle > 0$ that

$$\begin{aligned} \langle F_i(h(\boldsymbol{\omega} + \boldsymbol{\zeta}^*) + \mathbf{b}) - F_i(h(\boldsymbol{\omega}' + \boldsymbol{\zeta}^*) + \mathbf{b}) \rangle & \leq \widehat{\Delta} \langle h(\boldsymbol{\omega}' + \boldsymbol{\zeta}^*) + \mathbf{b} \rangle^{d-1} \langle h(\boldsymbol{\omega} - \boldsymbol{\omega}') \rangle \\ & \leq \widehat{\Delta} \langle h(\boldsymbol{\omega}' + \boldsymbol{\zeta}^*) \rangle^{d-1} \langle h(\boldsymbol{\omega} - \boldsymbol{\omega}') \rangle. \end{aligned}$$

Equation (6.5) gives $\langle \boldsymbol{\omega}' + \boldsymbol{\zeta}^* \rangle < \widehat{Q}$. This gives

$$\begin{aligned} \langle F_i(h(\boldsymbol{\omega} + \boldsymbol{\zeta}^*) + \mathbf{b}) - F_i(h(\boldsymbol{\omega}' + \boldsymbol{\zeta}^*) + \mathbf{b}) \rangle & \leq \widehat{\Delta} (\langle h \rangle^{d-1} (\widehat{Q} q^{-1})^{d-1}) \langle h \rangle \langle \boldsymbol{\omega} - \boldsymbol{\omega}' \rangle \\ & = \widehat{\Delta} \langle h \rangle^d \widehat{Q}^{d-1} q^{-(d-1)} \langle \boldsymbol{\omega} - \boldsymbol{\omega}' \rangle. \end{aligned}$$

Using the assumption (6.3) again, we get

$$\begin{aligned} & \langle \beta_i(F_i(h(\boldsymbol{\omega} + \boldsymbol{\zeta}^*) + \mathbf{b}) - F_i(h(\boldsymbol{\omega}' + \boldsymbol{\zeta}^*) + \mathbf{b})) \rangle \\ & \leq (\widehat{\Delta}^{-1} \langle h \rangle^{-d} q^{-1} \widehat{Q}^{1-d}) (\widehat{\Delta} \langle h \rangle^d \widehat{Q}^{d-1} q^{-(d-1)} \langle \boldsymbol{\omega} - \boldsymbol{\omega}' \rangle) \\ & = q^{-d} \langle \boldsymbol{\omega} - \boldsymbol{\omega}' \rangle, \end{aligned}$$

for each $i = 1, \dots, R$. Hence (6.20) follows as long as $\langle \boldsymbol{\omega} - \boldsymbol{\omega}' \rangle < q^{d-1}$. The integrand in (6.19) is indeed a continuous function of $\boldsymbol{\omega}$, when $\langle \boldsymbol{\omega} \rangle < \widehat{Q} \widehat{M}^{-1}$.

Hence we can apply Theorem A.1 in (6.19) with the substitution $\boldsymbol{\omega} = m\boldsymbol{\lambda}$. On recalling (6.5) and the assumption that $\langle m \rangle = \widehat{Q}$, this yields

$$U(\boldsymbol{\beta}) = \widehat{Q}^s \int_{\langle \boldsymbol{\lambda} \rangle < \widehat{M}^{-1}} E(\boldsymbol{\beta} \cdot \mathbf{F}(h(m\boldsymbol{\lambda} + \lfloor m\boldsymbol{\zeta} \rfloor_{Q-M}) + \mathbf{b})) \, d\boldsymbol{\lambda}.$$

Since \mathbf{F} is homogeneous of degree d , we can rewrite the above as

$$U(\boldsymbol{\beta}) = \widehat{Q}^s \int_{\boldsymbol{\lambda} < \widehat{M}^{-1}} E(m^d \boldsymbol{\beta} \cdot \mathbf{F}(h(\boldsymbol{\lambda} + \boldsymbol{\mu}) + \mathbf{c})) \, d\boldsymbol{\lambda}, \quad (6.21)$$

where

$$\boldsymbol{\mu} = m^{-1} \lfloor m\boldsymbol{\zeta} \rfloor_{Q-M} \quad \text{and} \quad \mathbf{c} = m^{-1} \mathbf{b}. \quad (6.22)$$

Note that

$$\boldsymbol{\mu} = m^{-1}(m\boldsymbol{\zeta} - \{m\boldsymbol{\zeta}\}_{Q-M}) = \boldsymbol{\zeta} - m^{-1} \{m\boldsymbol{\zeta}\}_{Q-M}.$$

Equation (6.4) and the assumption that $\langle m \rangle = \widehat{Q}$ imply that $\boldsymbol{\lambda} + \boldsymbol{\mu} \in \mathcal{E}$ whenever $\langle \boldsymbol{\lambda} \rangle < \widehat{M}^{-1}$. Putting $\boldsymbol{\sigma} = \boldsymbol{\lambda} + \boldsymbol{\mu}$ in (6.21), and using the translation-invariance of the measure, we obtain

$$U(\boldsymbol{\beta}) = \widehat{Q}^s \int_{\mathcal{E}} E(m^d \boldsymbol{\beta} \cdot \mathbf{F}(h\boldsymbol{\sigma} + \mathbf{c})) \, d\boldsymbol{\sigma}. \quad (6.23)$$

Our final step in proving (6.17) is to remove the \mathbf{c} in the integrand above. From (2.6) and (2.5), this is valid as long as

$$\langle m^d \boldsymbol{\beta} \cdot (\mathbf{F}(h\boldsymbol{\sigma} + \mathbf{c}) - \mathbf{F}(h\boldsymbol{\sigma})) \rangle < q^{-1}$$

for all $\boldsymbol{\sigma} \in \mathcal{E}$. Arguing as in the proof of (6.20), we see that when $i = 1, \dots, R$, we have

$$\begin{aligned} \langle m^d \beta_i(F_i(h\boldsymbol{\sigma} + \mathbf{c}) - F_i(h\boldsymbol{\sigma})) \rangle & = \langle m \rangle^d \langle \beta_i \rangle \langle F_i(h\boldsymbol{\sigma} + \mathbf{c}) - F_i(h\boldsymbol{\sigma}) \rangle \\ & \leq \langle m \rangle^d (\widehat{\Delta}^{-1} \widehat{Q}^{1-d} \langle h \rangle^{-d}) (\widehat{\Delta} \langle \mathbf{c} \rangle \langle h\boldsymbol{\sigma} \rangle^{d-1}). \end{aligned}$$

On recalling the definition of \mathbf{c} in (6.22), together with the assumptions that $\langle m \rangle = \widehat{Q}$, $\langle \mathbf{b} \rangle < \langle h \rangle$ and $\langle \boldsymbol{\sigma} \rangle < 1$, we get

$$\begin{aligned} \langle m^d \beta_i(F_i(h\boldsymbol{\sigma} + \mathbf{c}) - F_i(h\boldsymbol{\sigma})) \rangle & \leq \langle m \rangle^d (\widehat{\Delta}^{-1} \widehat{Q}^{1-d} \langle h \rangle^{-d}) (\widehat{\Delta} \langle m^{-1} \mathbf{b} \rangle \langle h \rangle^{d-1} \langle \boldsymbol{\sigma} \rangle^{d-1}) \\ & = (\langle \mathbf{b} \rangle \langle h \rangle^{-1}) \langle \boldsymbol{\sigma} \rangle^{d-1} \\ & \leq q^{-1} (q^{-1})^{d-1} < q^{-1} \end{aligned}$$

for all $i = 1, \dots, R$, as required. Removing the \mathbf{c} from the integral in (6.23), and once again using the homogeneity of the forms \mathbf{F} of degree d , we therefore obtain

$$U(\boldsymbol{\beta}) = \widehat{Q}^s \int_{\mathcal{E}} E(m^d \boldsymbol{\beta} \cdot \mathbf{F}(h\boldsymbol{\sigma})) \, d\boldsymbol{\sigma} = \widehat{Q}^s \int_{\mathcal{E}} E(h^d m^d \boldsymbol{\beta} \cdot \mathbf{F}(\boldsymbol{\sigma})) \, d\boldsymbol{\sigma}.$$

On recalling (6.2), this gives (6.17), as desired.

On substituting (6.17) into (6.16), and putting the resulting equation together with (6.15) into (6.14), the equality required by the lemma thus follows. \square

Recall (4.2), the choice of θ_0 in (5.5) and the assumption that $N \ll 1$. From these we deduce that when $\boldsymbol{\alpha} \in \mathfrak{N}(g, \mathbf{a})$, for some $\mathbf{a}/g \in \mathbb{K}^R$ for which the set $\mathfrak{N}(g, \mathbf{a})$ is defined, we have

$$\langle g \rangle < \widehat{P}^{R(d-1)\theta_0} = \widehat{P}^{1/2} < \widehat{P}\widehat{N}^{-1},$$

as long as P is large enough. Also, equations (4.2) and (5.5) give

$$\begin{aligned} \langle g\boldsymbol{\alpha} - \mathbf{a} \rangle &< \widehat{\Delta}^{-1} \langle h \rangle^{-d} \langle g \rangle \widehat{P}^{-d+R(d-1)\theta_0} < \widehat{\Delta}^{-1} \langle h \rangle^{-d} \widehat{P}^{-d+2R(d-1)\theta_0} \\ &= \widehat{\Delta}^{-1} \langle h \rangle^{-d} \widehat{P}^{1-d}. \end{aligned}$$

Hence every $\boldsymbol{\alpha} \in \mathfrak{N}(g, \mathbf{a})$ satisfies the conditions of Lemma 6.1 with $\mathcal{E} = \mathcal{B}$, $M = N$, $m = n$ and $Q = P$. Using that lemma provides the formula

$$T(\boldsymbol{\alpha}) = \widehat{P}^s \langle g \rangle^{-s} S(g, \mathbf{a}) I(h^d n^d \boldsymbol{\beta}; \mathcal{B}) \quad \text{for all } \boldsymbol{\alpha} \in \mathfrak{N}(g, \mathbf{a}), \quad (6.24)$$

where, as before, $\boldsymbol{\beta} = \boldsymbol{\alpha} - \mathbf{a}/g$.

For any monic $g \in \mathbb{A}$, write

$$\mathcal{A}(g) = \sum_{\substack{\langle \mathbf{a} \rangle < \langle g \rangle \\ (a_1, \dots, a_R, g)=1}} \langle g \rangle^{-s} S(g, \mathbf{a}). \quad (6.25)$$

When Y is a real number, set

$$\mathfrak{S}(Y) = \sum_{\langle g \rangle < \widehat{Y}}^{\dagger} \mathcal{A}(g) \quad (6.26)$$

and

$$J(Y) = \int_{\langle \boldsymbol{\beta} \rangle < \widehat{\Delta}^{-1} \langle h \rangle^{-d} \widehat{Y}} I(h^d n^d \boldsymbol{\beta}; \mathcal{B}) \, d\boldsymbol{\beta}. \quad (6.27)$$

Now integrate both sides of (6.24) with respect to all $\boldsymbol{\alpha} \in \mathfrak{N}(g, \mathbf{a})$, use the translation-invariance of the measure on \mathbb{T}^R , sum over all \mathbf{a} and g for which $\mathfrak{N}(g, \mathbf{a})$ is defined, and finally use Lemma 4.1 and the definitions (6.25), (6.26) and (6.27). By the choice of θ_0 in (5.5), we arrive at the formula

$$\begin{aligned} \rho_{h, \mathbf{b}}(n; \mathfrak{N}) &= \widehat{P}^s \mathfrak{S}(R(d-1)\theta_0 P) J((-d + R(d-1)\theta_0)P) \\ &= \widehat{P}^s \mathfrak{S}(P/2) J((-d + 1/2)P). \end{aligned} \quad (6.28)$$

With the asymptotic formula given in Theorem 1.1 in mind, the final aim of this section is to extract a factor of \widehat{P}^{-Rd} from the quantity $J((-d + 1/2)P)$. For any real number Y , let

$$\mathcal{J}(Y) = \int_{\langle \gamma \rangle < \widehat{\Delta}^{-1} \widehat{Y}} I(\gamma; \mathcal{B}) \, d\gamma. \quad (6.29)$$

We wish to make the substitution $\gamma = h^d n^d \beta$ in the integral

$$J((-d + 1/2)P) = \int_{\langle \beta \rangle < \widehat{\Delta}^{-1} \langle h \rangle^{-d} \widehat{P}^{-d+1/2}} I(h^d n^d \beta; \mathcal{B}) \, d\beta. \quad (6.30)$$

On recalling that $\langle n \rangle = \widehat{P}$, we see that this involves an application of Theorem A.1, which requires that the function $\gamma \mapsto I(\gamma; \mathcal{B})$ be continuous when $\langle \gamma \rangle < \widehat{\Delta}^{-1} \widehat{P}^{1/2}$. Again from (2.6) and (6.2), we see that this function takes a discrete set of values. So our requirement here is satisfied if we have

$$I(\gamma; \mathcal{B}) = I(\gamma'; \mathcal{B}) \quad \text{when } \langle \gamma - \gamma' \rangle \text{ is small enough.}$$

On recalling the definition (6.2) of $I(\gamma; \mathcal{B})$, it suffices to show that

$$E(\gamma \cdot \mathbf{F}(\sigma)) = E(\gamma' \cdot \mathbf{F}(\sigma)) \quad \text{for all } \sigma \in \mathcal{B}, \quad (6.31)$$

whenever $\langle \gamma - \gamma' \rangle$ is sufficiently small. From (2.6) and (2.5), we see that the displayed equality here holds once we have

$$\langle \gamma - \gamma' \rangle < (\max_{\sigma \in \mathcal{B}} \langle \mathbf{F}(\sigma) \rangle q)^{-1}.$$

From (2.11) and (2.2), the hypercube \mathcal{B} lies in the compact set \mathbb{T}^s . Thus the maximum above definitely exists. In conclusion, when γ and γ' satisfy the above inequality, the desired equality (6.31) holds, which implies that the function $\gamma \mapsto I(\gamma; \mathcal{B})$ is indeed continuous when $\langle \gamma \rangle < \widehat{\Delta}^{-1} \widehat{P}^{1/2}$.

Making the desired substitution $\gamma = h^d n^d \beta$ in (6.30), and using Theorem A.1 along with (6.29), we obtain

$$J((-d + 1/2)P) = \langle h \rangle^{-Rd} \widehat{P}^{-Rd} \mathcal{J}(P/2).$$

Putting this back into (6.28), we obtain the simplified formula

$$\rho_{h,b}(n; \mathfrak{N}) = \langle h \rangle^{-Rd} \widehat{P}^{s-Rd} \mathfrak{S}(P/2) \mathcal{J}(P/2). \quad (6.32)$$

7. SINGULAR SERIES

This section is devoted to the investigation of the factor $\mathfrak{S}(P/2)$ that appears in (6.32). To make the analysis of this factor more transparent, we extend the g -sum which defines it in (6.26) to the infinite sum

$$\mathfrak{S} = \sum_{g \in \mathbb{A}}^\dagger \mathcal{A}(g). \quad (7.1)$$

This quantity is known as the singular series associated to the system (2.10) of equations. We need to know whether the extension of sums alluded to above is valid, as well as the size of the error incurred in performing this extension. Moreover, to make the resulting asymptotic for $\rho_{h,b}(n; \mathfrak{N})$ useful,

we also require \mathfrak{S} to be positive. Equations (7.1) and (6.25) suggest that this information can be obtained by investigating the complete exponential sums $S(g, \mathbf{a})$. The first two issues can be resolved via an application of Lemma 3.7. On recalling (2.12) and (6.1), note that $S(g, \mathbf{a}) = T(\mathbf{a}/g; g, \mathbb{T}^s)$. That lemma therefore yields the upper bound

$$S(g, \mathbf{a}) \ll \langle g \rangle^s \left(\frac{\langle g \rangle^R}{\langle g \rangle^{Rd}} + \frac{1}{\langle g \rangle} \right)^{K/(R(d-1))} \ll \langle g \rangle^{s-K/(R(d-1))}.$$

Putting this into (6.25), we therefore obtain

$$\mathcal{A}(g) \ll \langle g \rangle^R \langle g \rangle^{-s} \langle g \rangle^{s-K/(R(d-1))} = \langle g \rangle^{R-K/(R(d-1))}. \quad (7.2)$$

Armed with this estimate, we are now able to confirm the convergence of \mathfrak{S} .

Lemma 7.1. *Under the assumptions (5.8) and $p > d$, we have*

$$\mathfrak{S} - \mathfrak{S}(Y) \ll \widehat{Y}^{R+1-K/(R(d-1))} \quad \text{for all real } Y.$$

In particular, the singular series \mathfrak{S} converges absolutely.

Proof. On recalling (6.26) and (7.1), and using (7.2), we obtain

$$\begin{aligned} \mathfrak{S} - \mathfrak{S}(Y) &= \sum_{\langle g \rangle \geq \widehat{Y}}^{\dagger} \mathcal{A}(g) \\ &\ll \sum_{\langle g \rangle \geq \widehat{Y}}^{\dagger} \langle g \rangle^{R-K/(R(d-1))} \\ &= \sum_{l \geq Y} (q^l)^{R-K/(R(d-1))} \sum_{\langle g \rangle = \widehat{l}}^{\dagger} 1 \\ &= \sum_{l \geq Y} (q^l)^{R+1-K/(R(d-1))} \\ &\ll \widehat{Y}^{R+1-K/(R(d-1))}. \end{aligned}$$

Here the last line follows from (5.8). This proves the first assertion of the lemma. The second assertion is an immediate consequence of this and (5.8). \square

It thus remains to establish the positivity of \mathfrak{S} . To this end, we first need the following standard factorisation property of the complete exponential sums $S(g, \mathbf{a})$.

Lemma 7.2. *Let $\mathbf{a}_1/g_1, \mathbf{a}_2/g_2 \in \mathbb{K}^R$ such that $(g_1, g_2) = 1$. Then*

$$S(g_1, \mathbf{a}_1)S(g_2, \mathbf{a}_2) = S(g_1g_2, g_1\mathbf{a}_2 + g_2\mathbf{a}_1).$$

Proof. Since $(g_1, g_2) = 1$, every s -tuple of residue classes $(\bmod g_1g_2)$ can be written uniquely as $g_1\mathbf{x}_2 + g_2\mathbf{x}_1$, where \mathbf{x}_u is a complete residue class $(\bmod$

g_u), for each $u = 1, 2$. Using this in (6.1) then yields

$$\begin{aligned} & S(g_1 g_2, g_1 \mathbf{a}_2 + g_2 \mathbf{a}_1) \\ &= \sum_{\langle \mathbf{x}_1 \rangle < \langle g_1 \rangle} \sum_{\langle \mathbf{x}_2 \rangle < \langle g_2 \rangle} E\left(\frac{(g_1 \mathbf{a}_2 + g_2 \mathbf{a}_1) \cdot \mathbf{F}(h(g_1 \mathbf{x}_2 + g_2 \mathbf{x}_1) + \mathbf{b})}{g_1 g_2}\right). \end{aligned}$$

Expanding out each polynomial $F_i(h(g_1 \mathbf{x}_2 + g_2 \mathbf{x}_1) + \mathbf{b})$ ($i = 1, \dots, R$), we see that

$$F_i(h(g_1 \mathbf{x}_2 + g_2 \mathbf{x}_1) + \mathbf{b}) \equiv \begin{cases} F_i(h(g_2 \mathbf{x}_1) + \mathbf{b}) & (\text{mod } g_1) \\ F_i(h(g_1 \mathbf{x}_2) + \mathbf{b}) & (\text{mod } g_2). \end{cases}$$

Hence

$$\begin{aligned} & S(g_1 g_2, g_1 \mathbf{a}_2 + g_2 \mathbf{a}_1) \\ &= \sum_{\langle \mathbf{x}_1 \rangle < \langle g_1 \rangle} \sum_{\langle \mathbf{x}_2 \rangle < \langle g_2 \rangle} E(\mathbf{a}_1 \cdot \mathbf{F}(h(g_2 \mathbf{x}_1) + \mathbf{b})/g_1) E(\mathbf{a}_2 \cdot \mathbf{F}(h(g_1 \mathbf{x}_2) + \mathbf{b})/g_2). \end{aligned}$$

Finally, since $(g_1, g_2) = 1$, as \mathbf{x}_1 (resp. \mathbf{x}_2) runs through all s -tuples of residue classes (mod g_1) (resp. mod g_2), so does $g_2 \mathbf{x}_1$ (resp. $g_1 \mathbf{x}_2$). On recalling (6.1), we obtain

$$\begin{aligned} & S(g_1 g_2, g_1 \mathbf{a}_2 + g_2 \mathbf{a}_1) \\ &= \left(\sum_{\langle \mathbf{x}_1 \rangle < \langle g_1 \rangle} E(\mathbf{a}_1 \cdot \mathbf{F}(h \mathbf{x}_1 + \mathbf{b})/g_1) \right) \left(\sum_{\langle \mathbf{x}_2 \rangle < \langle g_2 \rangle} E(\mathbf{a}_2 \cdot \mathbf{F}(h \mathbf{x}_2 + \mathbf{b})/g_2) \right) \\ &= S(g_1, \mathbf{a}_1) S(g_2, \mathbf{a}_2), \end{aligned}$$

as required. \square

For any ϖ , write

$$\Omega(\varpi) = 1 + \sum_{\nu=1}^{\infty} \mathcal{A}(\varpi^\nu). \quad (7.3)$$

Lemma 7.2 is then the ignition spark for the following Euler product representation of \mathfrak{S} .

Lemma 7.3. *We can write*

$$\mathfrak{S} = \prod_{\varpi} \Omega(\varpi).$$

Further, there exists a constant $Z = Z(s, R, d, \mathbf{F}, h, \mathbf{b})$ for which

$$\prod_{\langle \varpi \rangle \geq \widehat{Z}} \Omega(\varpi) \asymp 1.$$

Proof. On recalling the definition (6.25) of the arithmetic function \mathcal{A} , it follows from Lemma 7.2 that \mathcal{A} is multiplicative. The first assertion of this lemma thus follows by using this observation together with (7.3) on (7.1). To prove the second assertion, note that the upper bound part follows from the first

assertion together with Lemma 7.1. It remains to prove the lower bound part. An application of the upper bound (7.2) on (7.3) gives

$$\Omega(\varpi) - 1 = \sum_{\nu=1}^{\infty} \mathcal{A}(\varpi^\nu) \ll \sum_{\nu=1}^{\infty} \langle \varpi^\nu \rangle^{R-K/(R(d-1))} \ll \langle \varpi \rangle^{R-K/(R(d-1))}.$$

There thus exists a constant $Z = Z(s, d, R, \mathbf{F}, h, \mathbf{b})$ for which

$$\sum_{\langle \varpi \rangle \geq \widehat{Z}} |\Omega(\varpi) - 1| \ll \sum_{\langle \varpi \rangle \geq \widehat{Z}} \langle \varpi \rangle^{R-K/(R(d-1))}.$$

By the polynomial prime number theorem (see Theorem 2.2 of [13]), the number of monic irreducible polynomials in \mathbb{A} with degree l is $O(q^l/l)$. It thus follows from the last displayed inequality and (5.8) that

$$\begin{aligned} \sum_{\langle \varpi \rangle \geq \widehat{Z}} |\Omega(\varpi) - 1| &\ll \sum_{l \geq Z} (q^l)^{R-K/(R(d-1))} \sum_{\langle \varpi \rangle = l} 1 \\ &\ll \sum_{l \geq Z} (q^l)^{R+1-K/(R(d-1))} / l \ll 1. \end{aligned}$$

On extracting logarithms, we find that

$$\prod_{\langle \varpi \rangle \geq \widehat{Z}} \Omega(\varpi) \gg 1.$$

The lower bound part in the second assertion of the lemma duly follows. \square

Hence, to establish that $\mathfrak{S} > 0$, it suffices to show that $\Omega(\varpi) > 0$ for all ϖ with $\langle \varpi \rangle \ll 1$. There is a standard interpretation of $\Omega(\varpi)$ as the density of ϖ -adic solutions of the simultaneous equations in (2.10). To make this precise, we have to introduce more notation. For any monic $g \in \mathbb{A}$, let $M(g)$ denote the number of solutions of the simultaneous congruences

$$\mathbf{F}(h\mathbf{x} + \mathbf{b}) \equiv \mathbf{0} \pmod{g}. \quad (7.4)$$

We aim to relate $\Omega(\varpi)$ to $M(\varpi^k)$ for large integers k . The first step to achieving this is the following lemma concerning orthogonality relations for complete exponential sums.

Lemma 7.4. *Let $y, g \in \mathbb{A}$ with $g \neq 0$. Then*

$$\sum_{\langle x \rangle < \langle g \rangle} E(xy/g) = \begin{cases} \langle g \rangle, & \text{if } g \mid y, \\ 0, & \text{otherwise.} \end{cases}$$

Proof. This is Lemma 1(g) of [10]. \square

Our objective, alluded to before the last lemma, is made exact by the following lemma.

Lemma 7.5. *For any ϖ and positive integer k , we have*

$$M(\varpi^k) = \langle \varpi \rangle^{k(s-R)} \sum_{\nu=0}^k \mathcal{A}(\varpi^\nu).$$

In particular, we have

$$\Omega(\varpi) = \lim_{k \rightarrow \infty} \langle \varpi \rangle^{-k(s-R)} M(\varpi^k).$$

Proof. Applying Lemma 7.4 on the counting function $M(\varpi^k)$, we obtain

$$M(\varpi^k) = \langle \varpi \rangle^{-Rk} \sum_{\langle \mathbf{u} \rangle < \langle \varpi \rangle^k} \sum_{\langle \mathbf{x} \rangle < \langle \varpi \rangle^k} E(\mathbf{u} \cdot \mathbf{F}(h\mathbf{x} + \mathbf{b})/\varpi^k).$$

Now put $\mathbf{u} = \varpi^{k-\nu} \mathbf{w}$, where $\nu = 0, 1, \dots, k$ and $\mathbf{w} \in \mathbb{A}^R$ with the property that $(w_1, \dots, w_R, \varpi) = 1$. This gives

$$M(\varpi^k) = \langle \varpi \rangle^{-Rk} \sum_{\nu=0}^k \sum_{\substack{\langle \mathbf{w} \rangle < \langle \varpi \rangle^\nu \\ (w_1, \dots, w_R, \varpi)=1}} \sum_{\langle \mathbf{x} \rangle < \langle \varpi \rangle^k} E(\mathbf{w} \cdot \mathbf{F}(h\mathbf{x} + \mathbf{b})/\varpi^\nu).$$

The innermost sum here is periodic $(\bmod \varpi^\nu)$, hence

$$M(\varpi^k) = \langle \varpi \rangle^{-Rk} \sum_{\nu=0}^k \langle \varpi \rangle^{(k-\nu)s} \sum_{\substack{\langle \mathbf{w} \rangle < \langle \varpi \rangle^\nu \\ (w_1, \dots, w_R, \varpi)=1}} \sum_{\langle \mathbf{x} \rangle < \langle \varpi \rangle^\nu} E(\mathbf{w} \cdot \mathbf{F}(h\mathbf{x} + \mathbf{b})/\varpi^\nu).$$

On recalling (6.1) and (6.25), we obtain

$$\begin{aligned} M(\varpi^k) &= \langle \varpi \rangle^{k(s-R)} \sum_{\nu=0}^k \langle \varpi \rangle^{-\nu s} \sum_{\substack{\langle \mathbf{w} \rangle < \langle \varpi \rangle^\nu \\ (w_1, \dots, w_R, \varpi)=1}} S(\varpi^\nu, \mathbf{w}) \\ &= \langle \varpi \rangle^{k(s-R)} \sum_{\nu=0}^k \mathcal{A}(\varpi^\nu). \end{aligned}$$

This proves the first assertion of the lemma. The second assertion follows immediately from rearranging this equality, and recalling the definition (7.3) of $\Omega(\varpi)$. \square

The second assertion of the above lemma implies that, in order to have $\Omega(\varpi) > 0$, it suffices to show that

$$M(\varpi^k) \gg \langle \varpi \rangle^{k(s-R)} \quad \text{for all large } k, \quad (7.5)$$

where the implied constant here could depend on ϖ , but is independent of k . The plan is to show that, as long as the simultaneous equations (2.10) have a non-singular ϖ -adic solution for every ϖ , then the desired inequality (7.5) holds. When $\mathbf{x} \in \mathbb{K}_\infty^s$, write $\mathcal{F}(\mathbf{x})$ for the matrix with (i, j) -entry $(\partial/\partial x_j)(F_i(h\mathbf{x} + \mathbf{b}))$, for any $i = 1, \dots, R$ and $j = 1, \dots, s$. The following lemma brings us one step closer to the desired relation (7.5) via a Hensel lifting argument.

Lemma 7.6. *Suppose that for every ϖ , there exists a positive integer l along with an s -tuple $\mathbf{x} \pmod{\varpi^{2l-1}}$ for which*

$$\mathbf{F}(h\mathbf{x} + \mathbf{b}) \equiv \mathbf{0} \pmod{\varpi^{2l-1}} \quad (7.6)$$

and

$$\text{rank } \mathcal{F}(\mathbf{x}) < R \pmod{\varpi^{l-1}} \text{ but } \text{rank } \mathcal{F}(\mathbf{x}) = R \pmod{\varpi^l}. \quad (7.7)$$

Then for every positive integer ν , there exists at least $\langle \varpi \rangle^{\nu(s-R)}$ tuples $\mathbf{y} \pmod{\varpi^{2l-1+\nu}}$ such that

$$\mathbf{F}(h\mathbf{y} + \mathbf{b}) \equiv \mathbf{0} \pmod{\varpi^{2l-1+\nu}} \quad (7.8)$$

and

$$\text{rank } \mathcal{F}(\mathbf{y}) < R \pmod{\varpi^{l-1}} \text{ but } \text{rank } \mathcal{F}(\mathbf{y}) = R \pmod{\varpi^l}. \quad (7.9)$$

Proof. This proof is based on the argument which led to Hensel's lemma, also used in the proof of Lemma 17.1 of [4]. We proceed by induction on ν . When $\nu = 1$, let \mathbf{v} be an s -tuple $\pmod{\varpi}$ to be specified. By the multidimensional Taylor's theorem, we have

$$\mathbf{F}(h(\mathbf{x} + \varpi^l \mathbf{v}) + \mathbf{b}) \equiv \mathbf{F}(h\mathbf{x} + \mathbf{b}) + \nabla \mathbf{F}(h\mathbf{x} + \mathbf{b}) \cdot \varpi^l h\mathbf{v} \pmod{\varpi^{2l}}.$$

On recalling the definition of the matrix $\mathcal{F}(\mathbf{x})$, we can rewrite the above as

$$\mathbf{F}(h(\mathbf{x} + \varpi^l \mathbf{v}) + \mathbf{b}) \equiv \mathbf{F}(h\mathbf{x} + \mathbf{b}) + \mathcal{F}(\mathbf{x}) \cdot \varpi^l \mathbf{v} \pmod{\varpi^{2l}}. \quad (7.10)$$

We want \mathbf{v} to satisfy the system of congruences

$$\mathbf{F}(h(\mathbf{x} + \varpi^l \mathbf{v}) + \mathbf{b}) \equiv \mathbf{0} \pmod{\varpi^{2l}},$$

so that $\mathbf{y} = \mathbf{x} + \varpi^l \mathbf{v}$ is a solution of the congruences in (7.8) with $\nu = 1$. From (7.10), the last displayed congruence relation is equivalent to

$$\mathbf{F}(h\mathbf{x} + \mathbf{b}) + \mathcal{F}(\mathbf{x}) \cdot \varpi^l \mathbf{v} \equiv \mathbf{0} \pmod{\varpi^{2l}}.$$

The assumptions (7.6) and (7.7) allow us to divide through the whole congruence by ϖ^{2l-1} and hence rewrite it as

$$\varpi^{-(2l-1)} \mathbf{F}(h\mathbf{x} + \mathbf{b}) + \varpi^{-(l-1)} \mathcal{F}(\mathbf{x}) \cdot \mathbf{v} \equiv \mathbf{0} \pmod{\varpi}.$$

The latter half of (7.7) implies that the rank of the matrix $\varpi^{-(l-1)} \mathcal{F}(\mathbf{x}) \pmod{\varpi}$ is R . Hence the system of linear congruences above has at least $\langle \varpi \rangle^{s-R}$ solutions in $\mathbf{v} \pmod{\varpi}$. For each of these tuples \mathbf{v} , the tuple $\mathbf{y} = \mathbf{x} + \varpi^l \mathbf{v}$ satisfies both (7.8) and (7.9) in the current case for ν . This finishes the proof of the base case of our induction.

Suppose the induction hypothesis holds for some positive integer ν . This provides at least $\langle \varpi \rangle^{\nu(s-R)}$ solutions of (7.8) which satisfy (7.9) as well. Let \mathbf{y} be one of these solutions. Again let \mathbf{v} be an s -tuple $\pmod{\varpi}$ to be chosen. Using the multidimensional Taylor's theorem once more, we obtain

$$\begin{aligned} \mathbf{F}(h(\mathbf{y} + \varpi^{l+\nu} \mathbf{v}) + \mathbf{b}) &\equiv \mathbf{F}(h\mathbf{y} + \mathbf{b}) + \nabla \mathbf{F}(h\mathbf{y} + \mathbf{b}) \cdot \varpi^{l+\nu} h\mathbf{v} \\ &= \mathbf{F}(h\mathbf{y} + \mathbf{b}) + \mathcal{F}(\mathbf{y}) \cdot \varpi^{l+\nu} \mathbf{v} \pmod{\varpi^{2l+\nu}}. \end{aligned} \quad (7.11)$$

As in the case where $\nu = 1$, we plan to choose \mathbf{v} such that $\mathbf{F}(h(\mathbf{y} + \varpi^{l+\nu} \mathbf{v}) + \mathbf{b}) \equiv \mathbf{0} \pmod{\varpi^{2l+\nu}}$. Using (7.11), we can rewrite this desired relation as

$$\mathbf{F}(h\mathbf{y} + \mathbf{b}) + \varpi^{l+\nu} \mathcal{F}(\mathbf{y}) \cdot \mathbf{v} \equiv \mathbf{0} \pmod{\varpi^{2l+\nu}}.$$

The conditions (7.8) and (7.9) satisfied by \mathbf{y} allow us to divide through this congruence by $\varpi^{2l-1+\nu}$, and hence rewrite it as

$$\varpi^{-(2l-1+\nu)} \mathbf{F}(h\mathbf{y} + \mathbf{b}) + \varpi^{-(l-1)} \mathcal{F}(\mathbf{y}) \cdot \mathbf{v} \equiv \mathbf{0} \pmod{\varpi}.$$

The latter half of (7.9) implies, as before, that this system of congruences has at least $\langle \varpi \rangle^{s-R}$ solutions in $\mathbf{v} \pmod{\varpi}$. So altogether there are at least $\langle \varpi \rangle^{\nu(s-R)} \langle \varpi \rangle^{s-R} = \langle \varpi \rangle^{(\nu+1)(s-R)}$ tuples $\mathbf{z} = \mathbf{y} + \varpi^{l+\nu} \mathbf{v}$ that satisfy (7.8) and (7.9), with \mathbf{y} and ν replaced by \mathbf{z} and $\nu + 1$ respectively. This completes the proof of the inductive step. The lemma thus follows by induction on ν . \square

We are now ready to establish that $\mathfrak{S} > 0$.

Corollary 7.7. *Suppose that the affine algebraic set W has a non-singular ϖ -adic point for every ϖ . Then $\mathfrak{S} > 0$.*

Proof. Courtesy of (2.10), our assumption here implies the existence of a ϖ -adic s -tuple \mathbf{y} for which $\text{rank } \mathcal{F}(\mathbf{y}) = R$ over \mathbb{K}_{ϖ} . There thus exists a positive integer l together with an s -tuple $\mathbf{x} \pmod{\varpi^{2l-1}}$ for which the conditions (7.6) and (7.7) of Lemma 7.6 are satisfied. Recalling that $M(\varpi^{2l-1+\nu})$ counts the number of solutions of the congruences (7.4) with $g = \varpi^{2l-1+\nu}$, and using Lemma 7.6, we obtain

$$M(\varpi^{2l-1+\nu}) \geq \langle \varpi \rangle^{\nu(s-R)} = C \langle \varpi \rangle^{(2l-1+\nu)(s-R)},$$

where $C = \langle \varpi \rangle^{(2l-1)(R-s)}$. Evidently C is independent of ν . We have thus established the desired inequality (7.5), with $k = 2l - 1 + \nu$. Lemma 7.5 therefore gives $\Omega(\varpi) > 0$ for every ϖ with $\langle \varpi \rangle < \widehat{Z}$, where Z is given in Lemma 7.3. That lemma then leads to the conclusion that $\mathfrak{S} > 0$. \square

8. SINGULAR INTEGRAL

Similar to the way that the singular series is treated in the previous section, our aim now is to extend the finite integral $\mathcal{J}(P/2)$ in the formula (6.32) to the *complete singular integral*

$$\mathcal{J} = \int_{\mathbb{K}_{\infty}^R} I(\gamma; \mathcal{B}) \, d\gamma = \lim_{Y \rightarrow \infty} \mathcal{J}(Y). \quad (8.1)$$

We therefore need to know whether this integral converges absolutely, and the size of the error incurred by replacing $\mathcal{J}(P/2)$ by \mathcal{J} in (6.32). To make the resulting asymptotic formula for $\rho_{h,b}(n; \mathfrak{N})$ useful, we also require that $\mathcal{J} > 0$, under certain conditions on the geometry of X . The first two pieces of information are obtained via an estimate for the exponential integral $I(\gamma; \mathcal{B})$ given in (6.2). This is provided by the following lemma.

Lemma 8.1. *Let $\gamma \in \mathbb{K}_{\infty}^R$ and $M \in \mathbb{N}$. Let \mathcal{C} be any hypercube in \mathbb{T}^s with sidelength \widehat{M}^{-1} . Then provided that $p > d$, we have*

$$I(\gamma; \mathcal{C}) \ll \widehat{M}^{-s} (1 + \widehat{M}^{-d} \langle \gamma \rangle)^{-K/(R(d-1))}.$$

Proof. From (6.2), we see that when $\langle \gamma \rangle \leq \widehat{M}^d$, the estimate given by the lemma is trivial. Hence we may suppose that

$$\langle \gamma \rangle > \widehat{M}^d \geq 1. \quad (8.2)$$

Let Q be a large parameter to be chosen. The plan is to apply Lemma 6.1 to relate $I(\gamma; \mathcal{C})$ to the generating function

$$V(\alpha) = \sum_{\mathbf{x} \in t^Q \mathcal{C}} E(\alpha \cdot \mathbf{F}(\mathbf{x})),$$

for some α to be specified, and then use Lemma 3.7 to obtain an upper bound for the generating function.

Provided that Q satisfies the conditions

$$\widehat{Q} > \widehat{M} \text{ and } \langle \alpha \rangle < \widehat{\Delta}^{-1} \widehat{Q}^{1-d},$$

Lemma 6.1 is applicable with $\mathbf{a} = 0$, $g = 1$, $h = 1$, $m = t^Q$ and $\mathbf{b} = \mathbf{0}$. From (6.1), we have $S(1, \mathbf{0}) = 1$, so Lemma 6.1 yields

$$V(\alpha) = \widehat{Q}^s I(t^{Qd} \alpha; \mathcal{C}). \quad (8.3)$$

Choose $\alpha = t^{-Qd} \gamma$. The second condition on Q above is thus satisfied as long as $\widehat{Q} > \widehat{\Delta} \langle \gamma \rangle$. Meanwhile, write $\mathcal{E} = t^M \mathcal{C}$. Then \mathcal{E} is a hypercube in \mathbb{T}^s . An application of Lemma 3.7 with $L = K$, $h = g = 1$, $\mathbf{b} = \mathbf{a} = \mathbf{0}$ and Q replaced by $Q - M$, gives the estimate

$$\begin{aligned} V(\alpha) &= \sum_{\mathbf{x} \in t^{Q-M} \mathcal{E}} E(\alpha \cdot \mathbf{F}(\mathbf{x})) \\ &\ll (\widehat{Q} \widehat{M}^{-1})^s \left(\frac{(\widehat{Q} \widehat{M}^{-1})^R + (\widehat{Q} \widehat{M}^{-1})^{Rd} \langle \alpha \rangle^R}{(\widehat{Q} \widehat{M}^{-1})^{Rd}} + \frac{1}{1 + (\widehat{Q} \widehat{M}^{-1})^d \langle \alpha \rangle} \right)^{K/(R(d-1))} \\ &= \widehat{Q}^s \widehat{M}^{-s} \left((\widehat{Q} \widehat{M}^{-1})^{R(1-d)} + \widehat{Q}^{-Rd} \langle \gamma \rangle^R + \frac{1}{1 + \widehat{M}^{-d} \langle \gamma \rangle} \right)^{K/(R(d-1))} \\ &\ll \widehat{Q}^s \widehat{M}^{-s} \left((\widehat{Q} \widehat{M}^{-1})^{R(1-d)} + \widehat{Q}^{-Rd} \langle \gamma \rangle^R + \widehat{M}^d \langle \gamma \rangle^{-1} \right)^{K/(R(d-1))}. \end{aligned}$$

Here the last two lines follow from our choice of α and the assumption (8.2) respectively. On combining this with (8.3), we get

$$I(\gamma; \mathcal{C}) \ll \widehat{M}^{-s} \left((\widehat{Q} \widehat{M}^{-1})^{R(1-d)} + \widehat{Q}^{-Rd} \langle \gamma \rangle^R + \widehat{M}^d \langle \gamma \rangle^{-1} \right)^{K/(R(d-1))}.$$

Note that the desired upper bound for $I(\gamma; \mathcal{C})$ follows if the last term in the bracket above dominates. This is indeed the case if we choose Q so large that

$$\widehat{Q} > \max \left\{ \widehat{M}, \widehat{\Delta} \langle \gamma \rangle, \langle \gamma \rangle^{(R+1)/(Rd)} \widehat{M}^{-1/R}, \widehat{M}^{1-d/(R(d-1))} \langle \gamma \rangle^{1/(R(d-1))} \right\}.$$

This completes the proof of the lemma. \square

We are now in a position to resolve the convergence issue of the singular integral.

Corollary 8.2. *Let Y be any real number with $Y \geq dN + \Delta$. Then when $K > R^2(d-1)$, we have*

$$\mathcal{J} - \mathcal{J}(Y) \ll \widehat{Y}^{R-K/(R(d-1))}.$$

In particular, the singular integral \mathcal{J} converges absolutely. Further, we have $\mathcal{J} \ll 1$.

Proof. Recall from (2.11) that \mathcal{B} is a hypercube with sidelength \widehat{N}^{-1} , with N absolutely bounded. By Lemma 8.1 followed by R applications of (2.8), we get

$$\begin{aligned} \mathcal{J} - \mathcal{J}(Y) &\ll \int_{\langle \gamma \rangle \geq \widehat{\Delta}^{-1} \widehat{Y}} |I(\gamma; \mathcal{B})| \, d\gamma \\ &\ll \int_{\langle \gamma \rangle \geq \widehat{\Delta}^{-1} \widehat{Y}} \langle \gamma \rangle^{-K/(R(d-1))} \, d\gamma \\ &= \sum_{l \geq Y - \Delta} (q^l)^{-K/(R(d-1))} \int_{\langle \gamma \rangle = \widehat{l}} d\gamma \\ &\ll \sum_{l \geq Y - \Delta} (q^l)^{-K/(R(d-1))} (q^l)^R \\ &\ll \widehat{Y}^{R-K/(R(d-1))}. \end{aligned}$$

Here the last line follows from the assumption that $K > R^2(d-1)$. This proves the first assertion of the corollary. The second assertion is an immediate consequence of this. Meanwhile, a parallel argument to the one given above yields

$$\begin{aligned} \mathcal{J}(Y) &\ll \widehat{N}^{-s} \int_{\langle \gamma \rangle < \widehat{\Delta}^{-1} \widehat{Y}} (1 + \widehat{N}^{-d} \langle \gamma \rangle)^{-K/(R(d-1))} \, d\gamma \\ &\ll \int_{\langle \gamma \rangle < \widehat{N}^d} 1 \, d\gamma + \int_{\widehat{N}^d \leq \langle \gamma \rangle < \widehat{Y}} \langle \gamma \rangle^{-K/(R(d-1))} \, d\gamma \\ &= \widehat{N}^d + \sum_{dN \leq l < Y} (q^l)^{-K/(R(d-1))} \int_{\langle \gamma \rangle = \widehat{l}} d\gamma \\ &\ll 1 + \sum_{dN \leq l < Y} (q^l)^{R-K/(R(d-1))} \\ &\ll 1. \end{aligned}$$

This establishes the third assertion of the corollary. \square

It therefore remains to establish that $\mathcal{J} > 0$. On considering the relation (8.1), it suffices to show that $\mathcal{J}(Y) \gg 1$ for large Y . Recalling the definitions (6.29) and (6.2) of $\mathcal{J}(Y)$ and $I(\gamma; \mathcal{B})$ respectively, and applying Fubini's

theorem, we obtain

$$\begin{aligned}\mathcal{J}(Y) &= \int_{\langle \gamma \rangle < \widehat{\Delta}^{-1} \widehat{Y}} I(\gamma; \mathcal{B}) \, d\gamma \\ &= \int_{\langle \gamma \rangle < \widehat{\Delta}^{-1} \widehat{Y}} \int_{\mathcal{B}} E(\gamma \cdot \mathbf{F}(\boldsymbol{\sigma})) \, d\boldsymbol{\sigma} d\gamma \\ &= \int_{\mathcal{B}} \int_{\langle \gamma \rangle < \widehat{\Delta}^{-1} \widehat{Y}} E(\gamma \cdot \mathbf{F}(\boldsymbol{\sigma})) \, d\gamma d\boldsymbol{\sigma}.\end{aligned}$$

With R applications of Lemma 1(f) of [10], the γ -integral above gives

$$\mathcal{J}(Y) = \widehat{\Delta}^{-R} \widehat{Y}^R M_{\mathcal{B}}(Y), \quad (8.4)$$

where $M_{\mathcal{B}}(Y) = \text{Vol } \mathcal{M}_{\mathcal{B}}(Y)$, with

$$\mathcal{M}_{\mathcal{B}}(Y) = \left\{ \boldsymbol{\sigma} \in \mathcal{B} : \langle \mathbf{F}(\boldsymbol{\sigma}) \rangle < \widehat{\Delta} \widehat{Y}^{-1} \right\}. \quad (8.5)$$

To have the relation $\mathcal{J}(Y) \gg 1$, it is thus sufficient to prove that

$$\mathcal{M}_{\mathcal{B}}(Y) \gg \widehat{Y}^{-R} \quad \text{for large } Y. \quad (8.6)$$

To this end, write

$$\mathcal{M}_{\mathcal{B}} = \bigcap_{Y \geq \Delta} \mathcal{M}_{\mathcal{B}}(Y).$$

Equations (8.5) and (1.1) provide the relation

$$\mathcal{M}_{\mathcal{B}} = \{ \boldsymbol{\sigma} \in \mathcal{B} : \mathbf{F}(\boldsymbol{\sigma}) = \mathbf{0} \} = \mathcal{B} \cap X. \quad (8.7)$$

The following lemma provides a sufficient geometric condition under which (8.6) is true.

Lemma 8.3. *Provided that $\dim \mathcal{M}_{\mathcal{B}} \geq s - R$, we have $\mathcal{J} > 0$.*

Proof. We use the same argument used in the proof of Lemma 2 of [14]. Write $\mathcal{B}(Y)$ for the set of all points in \mathcal{B} which are at least a distance \widehat{Y}^{-1} away from the boundary of \mathcal{B} . Set $\mathcal{M}' = \mathcal{M}_{\mathcal{B}} \cap \mathcal{B}(Y)$. Choose \mathcal{M}'' to be a component of \mathcal{M}' with precise dimension $s - R$. Choose another subset $\mathcal{M}''' \subseteq \mathcal{M}''$ which can be parametrised by $s - R$ coordinates. Without loss of generality, we can assume these to be the first $s - R$ coordinates $\sigma_1, \dots, \sigma_{s-R}$. This means that we can find a non-empty open set $\mathcal{O} \subseteq \mathbb{T}^{s-R}$ together with a map $\mathbf{h} = (h_1, \dots, h_R) : \mathcal{O} \rightarrow \mathbb{T}^R$ such that

$$(\sigma_1, \dots, \sigma_{s-R}, \mathbf{h}(\sigma_1, \dots, \sigma_{s-R})) \in \mathcal{M}''' \quad \text{when } (\sigma_1, \dots, \sigma_{s-R}) \in \mathcal{O}. \quad (8.8)$$

Since $\mathcal{M}''' \subseteq \mathcal{M}_{\mathcal{B}}$, equation (8.7) gives

$$\mathbf{F}(\sigma_1, \dots, \sigma_{s-R}, \mathbf{h}(\sigma_1, \dots, \sigma_{s-R})) = \mathbf{0} \quad \text{when } (\sigma_1, \dots, \sigma_{s-R}) \in \mathcal{O}. \quad (8.9)$$

For any positive integer T , write S_T for the set of s -tuples $(\sigma_1, \dots, \sigma_s) \in \mathbb{T}^s$ with $(\sigma_1, \dots, \sigma_{s-R}) \in \mathcal{O}$ such that

$$\langle \sigma_{s-R+i} - h_i(\sigma_1, \dots, \sigma_{s-R}) \rangle < \widehat{T}^{-1} \text{ for all } i = 1, \dots, R.$$

Our choice of \mathcal{M}''' indicates that all points of \mathcal{M}''' are in \mathcal{B} and at least a distance \widehat{Y}^{-1} away from the boundary of \mathcal{B} . It thus follows from (8.8) and the inequality above that $S_T \subseteq \mathcal{B}$ whenever $T > Y$. Further, observe that

$$\begin{aligned} & \langle (\sigma_1, \dots, \sigma_{s-R}, \mathbf{h}(\sigma_1, \dots, \sigma_{s-R})) - (\sigma_1, \dots, \sigma_s) \rangle \\ &= \max_{1 \leq i \leq R} \langle \sigma_{s-R+i} - h_i(\sigma_1, \dots, \sigma_{s-R}) \rangle. \end{aligned}$$

On recalling (8.9) and the fact that the forms \mathbf{F} have degree d with coefficients having absolute value at most $\widehat{\Delta}$, we deduce that when $(\sigma_1, \dots, \sigma_{s-R}) \in \mathcal{O}$, we have

$$\begin{aligned} \langle \mathbf{F}(\sigma_1, \dots, \sigma_s) \rangle &= \langle \mathbf{F}(\sigma_1, \dots, \sigma_s) - \mathbf{F}(\sigma_1, \dots, \sigma_{s-R}, \mathbf{h}(\sigma_1, \dots, \sigma_{s-R})) \rangle \\ &\leq \widehat{\Delta} \langle (\sigma_1, \dots, \sigma_s) \rangle^{d-1} \max_{1 \leq i \leq R} \langle \sigma_{s-R+i} - h_i(\sigma_1, \dots, \sigma_{s-R}) \rangle. \end{aligned}$$

In particular, when $(\sigma_1, \dots, \sigma_s) \in S_Y$, we get

$$\langle \mathbf{F}(\sigma_1, \dots, \sigma_s) \rangle < \widehat{\Delta} \widehat{Y}^{-1}.$$

From (8.5), the above inequality implies that $S_Y \subseteq \mathcal{M}_{\mathcal{B}}(Y)$. It is apparent from the definition of S_Y that $\text{Vol } S_Y \gg \widehat{Y}^{-R}$. The last inclusion relation thus implies that $\text{Vol } \mathcal{M}_{\mathcal{B}}(Y) \gg \widehat{Y}^{-R}$. This is exactly (8.6). The conclusion of the lemma thus follows by inserting this estimate into (8.4), and recalling the relation (8.1). \square

With this result, we only have to show that $\dim \mathcal{M}_{\mathcal{B}} \geq s - R$. But we have chosen our box \mathcal{B} to be centred at a non-singular real point on X . By the Implicit Function Theorem in arbitrary fields (see Theorem 2.2.1 in [9]), the set $\mathcal{M}_{\mathcal{B}}$ of real points in $X \cap \mathcal{B}$ has dimension at least $s - R$, as required. On combining this with Lemma 8.3, we conclude that $\mathcal{J} > 0$, under the condition that X contains a non-singular real point.

9. CONCLUSION

We are now in a position to settle the outstanding theorems in this paper. We first derive an asymptotic formula for the number $\rho_{h,\mathbf{b}}(n)$ of integral solutions \mathbf{x} to (2.10) with $\mathbf{x} \in n\mathcal{B}_N$. Applying Lemma 7.1 and Corollary 8.2 in (6.32) with $Y = P/2$, we obtain

$$\rho_{h,\mathbf{b}}(n; \mathfrak{N}) = \langle h \rangle^{-Rd} \widehat{P}^{s-Rd} (\mathfrak{S}\mathcal{J} + O((\widehat{P}^{1/2})^{R+1-K/(R(d-1))})).$$

On recalling that $\mathbb{T}^R = \mathfrak{N} \cup \mathfrak{n}$, and using Lemma 5.3, we obtain

$$\begin{aligned} \rho_{h,\mathbf{b}}(n) &= \rho_{h,\mathbf{b}}(n; \mathfrak{N}) + \rho_{h,\mathbf{b}}(n; \mathfrak{n}) \\ &= \langle h \rangle^{-Rd} \widehat{P}^{s-Rd} (\mathfrak{S}\mathcal{J} + O(\widehat{P}^{-(K/(R(d-1))-R-1)/4})), \end{aligned} \quad (9.1)$$

for every monic polynomials $n, h \in \mathbb{A}$ with $\langle n \rangle = \widehat{P}$, and every integral s -tuple \mathbf{b} with $\langle \mathbf{b} \rangle < \langle h \rangle$.

Now we make different choices for h, \mathbf{b}, n and N in deriving Theorems 1.1 and 1.2. To establish Theorem 1.1, we take $h = 1, \mathbf{b} = \mathbf{0}, n = t^P, N = 0$ and $\boldsymbol{\xi}$ to be an arbitrary point in \mathbb{T}^s . From (2.11), we see that $\mathcal{B}_0 = \mathbb{T}^s$.

Meanwhile, on considering the respective underlying equations (2.10) and (1.1) of the algebraic sets W and X respectively, we have $W = X$. Thus $\rho_{1,0}(n) = \rho(P)$, where $\rho(P)$ is defined just before Theorem 1.1. On recalling (1.4) and (5.1), equation (9.1) thus gives

$$\rho(P) = \widehat{P}^{s-Rd}(\mathfrak{S}\mathcal{J} + O(\widehat{P}^{-k/(2^{d+1}R(d-1))})).$$

This is the asymptotic formula required in Theorem 1.1. Further, Corollary 7.7 and the discussion after Lemma 8.3 imply that the quantities \mathfrak{S} and \mathcal{J} are indeed positive under the hypotheses of Theorem 1.1. The proof of that theorem is thus complete.

Before proving Theorem 1.2, we introduce a few more notations. For every ϖ , let $\langle \cdot \rangle_{\varpi}$ be the ϖ -adic absolute value on \mathbb{K} . More precisely, when $x \in \mathbb{A} \setminus \{0\}$, we put $\langle x \rangle_{\varpi} = \langle \varpi \rangle^{-l}$, where l is the non-negative integer such that $\varpi^l \| x$. By convention, put $\langle 0 \rangle_{\varpi} = 0$. We extend this absolute value to \mathbb{K} by the relation

$$\langle x/y \rangle_{\varpi} = \langle x \rangle_{\varpi} / \langle y \rangle_{\varpi} \quad (x, y \in \mathbb{A}, y \neq 0).$$

Let \mathbb{A}_{ϖ} and \mathbb{K}_{ϖ} denote respectively the completions of \mathbb{A} and \mathbb{K} with respect to this absolute value.

The proof of Theorem 1.2 is largely based on the argument which led to Corollary 1 of [16]. Note that (9.1) has the following immediate consequence.

Lemma 9.1. *Suppose \mathbf{F} has a non-singular real zero $\boldsymbol{\xi}$. Let $h \in \mathbb{A}$, $\mathbf{b} \in \mathbb{A}^s$, with h monic and $\langle \mathbf{b} \rangle < \langle h \rangle$, such that for every ϖ , there exists a non-singular solution \mathbf{x} of the equations $\mathbf{F}(h\mathbf{x} + \mathbf{b}) = \mathbf{0}$, with $\mathbf{x} \in \mathbb{A}_{\varpi}^s$. Then for every positive integer N , there exists a $\boldsymbol{\xi}_0 \in \mathbb{A}^s$ such that*

$$\mathbf{F}(h\boldsymbol{\xi}_0 + \mathbf{b}) = \mathbf{0}, \tag{9.2}$$

and there exists a positive constant $C = C(s, d, R, \mathbf{F}, h, \mathbf{b}, N)$ such that whenever n is a monic polynomial in \mathbb{A} with $\langle n \rangle \geq \widehat{C}$, we have

$$\langle n^{-1}\boldsymbol{\xi}_0 - \boldsymbol{\xi} \rangle < \widehat{N}^{-1}. \tag{9.3}$$

Proof. From (2.11) and the definition of $\rho_{h,\mathbf{b}}(n)$ given before (2.14), we see that $\rho_{h,\mathbf{b}}(n)$ counts the number of integral points $\boldsymbol{\xi}_0$ which satisfy (9.2) and (9.3). Hence, under the conditions that $\mathfrak{S} > 0$ and $\mathcal{J} > 0$, the existence of such $\boldsymbol{\xi}_0$ is confirmed by (9.1) when $\langle n \rangle = \widehat{P} \geq \widehat{C}$, for some positive constant $C = C(s, d, R, \mathbf{F}, h, \mathbf{b}, N)$. But our present hypothesis on the existence of non-singular ϖ -adic solutions for every ϖ renders Corollary 7.7 applicable. That corollary gives $\mathfrak{S} > 0$. Also, now that X has a non-singular real point, the conclusion of §8 implies that $\mathcal{J} > 0$. This completes the proof of the lemma. \square

We now apply the above lemma to establish Theorem 1.2, which we reformulate below.

Lemma 9.2. *Suppose the hypotheses of Theorem 1.2 hold. Let $\boldsymbol{\zeta}_{\infty} \in X(\mathbb{K}_{\infty})$. Fix a collection of distinct monic irreducible polynomials $\varpi_1, \dots, \varpi_r$. For each*

$i = 1, \dots, r$, let $\zeta_i \in X(\mathbb{K}_{\varpi_i})$. Let $N_\infty, N_1, \dots, N_r$ be positive integers. Then there exists a $\zeta \in X(\mathbb{K})$ such that

$$\langle \zeta - \zeta_\infty \rangle < \widehat{N}_\infty^{-1}$$

and

$$\langle \zeta - \zeta_i \rangle_{\varpi_i} \leq \langle \varpi_i \rangle^{-N_i} \quad \text{for all } i = 1, \dots, r.$$

Proof. Let $h = \varpi_1^{N_1} \dots \varpi_r^{N_r}$ and $\xi = h^{-1} \zeta_\infty$. By the homogeneity of \mathbf{F} , we can assume without loss of generality that $\zeta_i \in \mathbb{A}_{\varpi_i}^s$ for each $i = 1, \dots, r$. Let $\mathbf{b}_i \in \mathbb{A}^s$ such that $\zeta_i \equiv \mathbf{b}_i \pmod{\varpi_i^{N_i}}$. By the Chinese Remainder Theorem, we can find a $\mathbf{b} \in \mathbb{A}^s$ such that $\mathbf{b} \equiv \mathbf{b}_i \pmod{\varpi_i^{N_i}}$, for each i .

We first verify that, with the above choices of h and \mathbf{b} , the ϖ -adic solubility condition of Lemma 9.1 is satisfied for every ϖ . First consider the case where $\varpi \notin \{\varpi_1, \dots, \varpi_r\}$. The hypotheses of Theorem 1.2 imply the existence of a $\zeta^{(\varpi)} \in X(\mathbb{K}_\varpi)$. By homogeneity again, we can assume that $\zeta^{(\varpi)} \in X(\mathbb{A}_\varpi)$. Take $\mathbf{x}^{(\varpi)} = (\zeta^{(\varpi)} - \mathbf{b})/h$. Then $\mathbf{F}(h\mathbf{x}^{(\varpi)} + \mathbf{b}) = \mathbf{0}$. Since $(h, \varpi) = 1$, it follows that $\mathbf{x}^{(\varpi)} \in \mathbb{A}_\varpi^s$. Further, now that X is non-singular, it follows that $\mathbf{x}^{(\varpi)}$ is a non-singular solution of the equations $\mathbf{F}(h\mathbf{x} + \mathbf{b}) = \mathbf{0}$. Now consider the monic irreducibles amongst $\varpi_1, \dots, \varpi_r$. Fix $i \in \{1, \dots, r\}$. As in the previous case, take $\mathbf{x}_i = (\zeta_i - \mathbf{b})/h$. Since $\varpi_i^{N_i} \parallel h$ and $\zeta_i \equiv \mathbf{b} \pmod{\varpi_i^{N_i}}$, it follows that $\mathbf{x}_i \in \mathbb{A}_{\varpi_i}^s$. The non-singularity of X once again implies that \mathbf{x}_i is a non-singular solution of $\mathbf{F}(h\mathbf{x} + \mathbf{b}) = \mathbf{0}$. We have thus shown that for every ϖ , the ϖ -adic solubility condition of Lemma 9.1 is indeed satisfied.

We apply that lemma with $\widehat{N} = \widehat{N}_\infty \langle h \rangle$. This gives an integral s -tuple ξ_0 such that

$$\mathbf{F}(h\xi_0 + \mathbf{b}) = \mathbf{0}, \tag{9.4}$$

together with a positive constant $C = C(s, d, R, \mathbf{F}, h, \mathbf{b})$ such that whenever n is a monic polynomial in \mathbb{A} with $\langle n \rangle \geq \widehat{C}$, we have

$$\langle n^{-1}\xi_0 - \xi \rangle < \widehat{N}_\infty^{-1} \langle h \rangle^{-1}.$$

Recalling our choice of ξ at the beginning of this proof, we multiply through the last relation by h and obtain

$$\langle hn^{-1}\xi_0 - \zeta_\infty \rangle < \widehat{N}_\infty^{-1}. \tag{9.5}$$

Choose n to be a monic polynomial with

$$n \equiv 1 \pmod{\varpi_1^{N_1} \dots \varpi_r^{N_r}}, \tag{9.6}$$

such that

$$\langle n \rangle > \max \left\{ \langle \mathbf{b} \rangle \widehat{N}_\infty, \widehat{C} \right\}. \tag{9.7}$$

Take $\zeta = n^{-1}(h\xi_0 + \mathbf{b})$. By the homogeneity of \mathbf{F} together with (9.4), we have $\zeta \in X(\mathbb{K})$. Using (9.5) together with (9.7), we also get

$$\begin{aligned} \langle \zeta - \zeta_\infty \rangle &= \langle n^{-1}(h\xi_0 + \mathbf{b}) - \zeta_\infty \rangle \\ &\leq \max \left\{ \langle hn^{-1}\xi_0 - \zeta_\infty \rangle, \langle \mathbf{b} \rangle / \langle n \rangle \right\} < \widehat{N}_\infty^{-1}. \end{aligned}$$

Finally, our choice of ζ implies that

$$\begin{aligned} \langle \zeta - \zeta_i \rangle_{\varpi_i} &\leq \max \{ \langle \zeta - \zeta_i/n \rangle_{\varpi_i}, \langle \zeta_i/n - \zeta_i \rangle_{\varpi_i} \} \\ &= \max \{ \langle h\xi_0 + \mathbf{b} - \zeta_i \rangle_{\varpi_i} / \langle n \rangle_{\varpi_i}, \langle \zeta_i \rangle_{\varpi_i} \langle n-1 \rangle_{\varpi_i} / \langle n \rangle_{\varpi_i} \}. \end{aligned} \quad (9.8)$$

Moreover, our choices of h and \mathbf{b} reveal that $h \equiv 0 \pmod{\varpi_i^{N_i}}$ and $\mathbf{b} \equiv \zeta_i \pmod{\varpi_i^{N_i}}$. Hence

$$\langle h\xi_0 + \mathbf{b} - \zeta_i \rangle_{\varpi_i} \leq \langle \varpi_i \rangle^{-N_i}.$$

Meanwhile, relation (9.6) gives $\langle n-1 \rangle_{\varpi_i} \leq \langle \varpi_i \rangle^{-N_i}$ and $\langle n \rangle_{\varpi_i} = 1$. These observations together with (9.8) thus imply that

$$\langle \zeta - \zeta_i \rangle_{\varpi_i} \leq \langle \varpi_i \rangle^{-N_i} \quad \text{for all } i = 1, \dots, r.$$

This completes the proof of the lemma, and thus that of Theorem 1.2. \square

Our task in this paper is therefore complete, subject to the resolution of a number of technical results given in the two appendices below.

APPENDIX A. INTEGRATION OVER FUNCTION FIELDS

In this appendix, we state and prove the change-of-variable property for integrals over function fields. We investigate integrals in the form

$$\int_{\Gamma} f(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha},$$

for some suitable function $f : \Gamma \rightarrow \mathbb{C}$ defined on some D -dimensional box Γ . Henceforth, a D -dimensional box refers to the set of real D -tuples \mathbf{x} satisfying constraints of the form

$$\langle x_u \rangle < \widehat{R}_u \quad (u = 1, \dots, D), \quad (\text{A.1})$$

for some real numbers R_u ($1 \leq u \leq D$). The special case where $D = 1$ and $f = E \circ \varphi$, where E is defined in (2.6) and φ is a polynomial map in one variable, is investigated in Proposition I.6 of [3]. The aim here is then to extend that result to rather more general functions defined over a box as in (A.1).

For convenience of exposition, we introduce further notations. For any $\zeta \in \mathbb{K}_{\infty} \setminus \{0\}$ and any integer Y , write $\Theta_Y(\zeta)$ for the coefficient of t^Y in the expansion for ζ . By convention, we define $\Theta_Y(0) = 0$ for any integer Y . In addition, for every integer L , write \mathcal{D}_L for the set of all $\alpha \in \mathbb{K}_{\infty}$ such that $\langle \alpha \rangle < \widehat{L}$. We ultimately establish the following theorem.

Theorem A.1. *Let D be a positive integer, and let Γ be any box in \mathbb{K}_{∞}^D . Let $f : \Gamma \rightarrow \mathbb{C}$ be continuous. Suppose $M \in GL_D(\mathbb{K}_{\infty})$. Then*

$$\int_{\Gamma} f(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} = \langle \det M \rangle \int_{M\Gamma} f(M\boldsymbol{\gamma}) \, d\boldsymbol{\gamma}. \quad (\text{A.2})$$

Since $GL_D(\mathbb{K}_\infty)$ is generated by the elementary matrices, namely permutation, diagonal and unipotent matrices, it suffices to prove Theorem A.1 when M is one of these matrices. When M is a permutation matrix, Theorem A.1 is equivalent to the property that we can relabel the variables of integration without changing the value of the integral of interest. This is evidently true, so we concentrate on the remaining two cases for M . Induction on the dimension D is central to our treatment of these cases. Hence, for diagonal matrices M , we first require the following lemma, which deals with the inductive base where $D = 1$.

Lemma A.2. *Let $f : \mathcal{D}_L \rightarrow \mathbb{C}$ be a continuous function. Let L be an integer, and $\beta \in \mathbb{K}_\infty \setminus \{0\}$. Then*

$$\int_{\langle \alpha \rangle < \widehat{L}} f(\alpha) \, d\alpha = \langle \beta \rangle \int_{\langle \gamma \rangle < \widehat{L} \langle \beta \rangle^{-1}} f(\beta \gamma) \, d\gamma. \quad (\text{A.3})$$

Proof. Since f is continuous on the compact set \mathcal{D}_L , it is uniformly continuous. For the sake of simplicity, when $H, J \in \mathbb{Z}$ with $-H \leq J-1$, and $a_{-H}, \dots, a_{J-1} \in \mathbb{F}_q$, write

$$g = g(\mathbf{a}; H, J) = a_{-H}t^{-H} + \dots + a_{J-1}t^{J-1}.$$

First we prove that for any integer J and any uniformly continuous function h on \mathcal{D}_J , we have

$$\int_{\langle \alpha \rangle < \widehat{J}} h(\alpha) \, d\alpha = \lim_{H \rightarrow \infty} \lambda(H, J), \quad (\text{A.4})$$

where

$$\lambda(H, J) = \widehat{H}^{-1} \sum_{\substack{a_k \in \mathbb{F}_q \\ (-H \leq k \leq J-1)}} h(g). \quad (\text{A.5})$$

For any $\alpha \in \mathcal{D}_J$, we make use of the substitution $\alpha = g + \gamma$, where the a_k ($-H \leq k \leq J-1$) run through elements of \mathbb{F}_q and γ runs through elements of \mathbb{K}_∞ with $\langle \gamma \rangle < \widehat{H}^{-1}$. This yields

$$\int_{\langle \alpha \rangle < \widehat{J}} h(\alpha) \, d\alpha = \int_{\langle \gamma \rangle < \widehat{H}^{-1}} \sum_{\substack{a_k \in \mathbb{F}_q \\ (-H \leq k \leq J-1)}} h(g + \gamma) \, d\gamma,$$

for all positive integers H . Meanwhile, on recalling (2.8), equation (A.5) can be rewritten as

$$\lambda(H, J) = \int_{\langle \gamma \rangle < \widehat{H}^{-1}} \sum_{\substack{a_k \in \mathbb{F}_q \\ (-H \leq k \leq J-1)}} h(g) \, d\gamma.$$

To prove (A.4), it thus suffices to show that, for each fixed choice of J and any $\varepsilon > 0$, there exists a positive integer $H_0 = H_0(\varepsilon, J)$ such that when $H \in \mathbb{Z}$ with $H \geq H_0$, we have

$$\left| \int_{\langle \gamma \rangle < \widehat{H}^{-1}} \sum_{\substack{a_k \in \mathbb{F}_q \\ (-H \leq k \leq J-1)}} h(g + \gamma) \, d\gamma - \int_{\langle \gamma \rangle < \widehat{H}^{-1}} \sum_{\substack{a_k \in \mathbb{F}_q \\ (-H \leq k \leq J-1)}} h(g) \, d\gamma \right| < \varepsilon.$$

There are exactly $q^{J-(-H)} = \widehat{J}\widehat{H}$ choices for the a_k ($-H \leq k \leq J-1$) in the sum above, whereas the γ -integral above is over a set of measure \widehat{H}^{-1} , courtesy of (2.8). We can therefore reduce the task of establishing the last displayed relation to showing that, for every $\varepsilon > 0$, there exists a positive integer $H_0 = H_0(\varepsilon, J)$ such that whenever $H \geq H_0$, we have

$$|h(g + \gamma) - h(g)| < \varepsilon \widehat{J}^{-1},$$

for any $\gamma \in \mathcal{D}_{-H}$ and any $a_k \in \mathbb{F}_q$ ($-H \leq k \leq J-1$). But this is simply a consequence of the assumption that h is uniformly continuous.

Write $\widehat{M} = \langle \beta \rangle$, for some integer M . Using (A.4) with $h = f$, $J = L$ and $H = G - M$, we can rewrite the left side of (A.3) as

$$\begin{aligned} & \int_{\langle \alpha \rangle < \widehat{L}} f(\alpha) \, d\alpha \\ &= \widehat{M} \lim_{G \rightarrow \infty} \widehat{G}^{-1} \sum_{\substack{a_k \in \mathbb{F}_q \\ (-G+M \leq k \leq L-1)}} f(a_{-G+M}t^{-G+M} + \dots + a_{L-1}t^{L-1}). \end{aligned}$$

When $c_j \in \mathbb{F}_q$ ($j \leq L - M - 1$), put

$$w = w(\mathbf{c}; L, M) = \sum_{j=-\infty}^{L-M-1} c_j t^j. \quad (\text{A.6})$$

In terms of the truncation operation defined in (2.3), we have

$$c_{-G}t^{-G} + \dots + c_{L-M-1}t^{L-M-1} = [w]_{-G}.$$

Put $v(\gamma) = f(\beta\gamma)$. Evidently v is still a continuous function on \mathcal{D}_{L-M} . Using (A.4) again with $H = G$, $h = v$ and $\widehat{J} = \widehat{L}\widehat{M}^{-1}$, we can rewrite the right side of (A.3) as

$$\begin{aligned} & \langle \beta \rangle \int_{\langle \gamma \rangle < \widehat{L}\langle \beta \rangle^{-1}} f(\beta\gamma) \, d\gamma \\ &= \widehat{M} \lim_{G \rightarrow \infty} \widehat{G}^{-1} \sum_{\substack{c_r \in \mathbb{F}_q \\ (-G \leq r \leq L-M-1)}} v(c_{-G}t^{-G} + \dots + c_{L-M-1}t^{L-M-1}) \\ &= \widehat{M} \lim_{G \rightarrow \infty} \widehat{G}^{-1} \sum_{\substack{c_r \in \mathbb{F}_q \\ (-G \leq r \leq L-M-1)}} f(\beta[w]_{-G}). \end{aligned}$$

To prove the lemma, it thus remains to show that, when L and β are fixed as in the statement of the lemma, and G is sufficiently large in terms of L and M , we have

$$\begin{aligned} & \sum_{\substack{a_k \in \mathbb{F}_q \\ (-G+M \leq k \leq L-1)}} f(a_{-G+M}t^{-G+M} + \dots + a_{L-1}t^{L-1}) \\ &= \sum_{\substack{c_r \in \mathbb{F}_q \\ (-G \leq r \leq L-M-1)}} f(\beta[w]_{-G}) + o(\widehat{G}). \end{aligned} \quad (\text{A.7})$$

To this end, let G be a large integer to be specified. We show that we can replace $f(\beta[w]_{-G})$ by $f(\lfloor \beta w \rfloor_{-G+M})$, allowing for a negligible error in the process. This allows us to work with the finite expansion $\lfloor \beta w \rfloor_{-G+M}$ in t in place of the infinite expansion $\beta[w]_{-G}$. This is valid if we can prove that, for any $\varepsilon > 0$, there exists a $G_0 = G_0(L, M, \varepsilon) > 0$ such that whenever $G \geq G_0$ and $w \in \mathcal{D}_{L-M}$, we have

$$|f(\beta[w]_{-G}) - f(\lfloor \beta w \rfloor_{-G+M})| < \varepsilon. \quad (\text{A.8})$$

But this follows from the uniform continuity of f . Indeed, note that

$$\begin{aligned} \beta[w]_{-G} - \lfloor \beta w \rfloor_{-G+M} &= (\beta w - \beta \{w\}_{-G}) - (\beta w - \{\beta w\}_{-G+M}) \\ &= \{\beta w\}_{-G+M} - \beta \{w\}_{-G}. \end{aligned}$$

On recalling that $\langle \beta \rangle = \widehat{M}$, we thus have

$$\langle \beta[w]_{-G} - \lfloor \beta w \rfloor_{-G+M} \rangle < q^{-G+M} = \widehat{G}^{-1} \widehat{M},$$

which can be made arbitrarily small by taking G to be large. This together with the uniform continuity of f implies (A.8), as a result. Since there are exactly $q^{L-M+G} = \widehat{L} \widehat{M}^{-1} \widehat{G}$ elements $c_r \in \mathbb{F}_q$ ($-G \leq r \leq L-M-1$), relation (A.8), with ε replaced by $\varepsilon \widehat{L}^{-1} \widehat{M}$ further implies that when G is large in terms of L and M , we have

$$\sum_{\substack{c_r \in \mathbb{F}_q \\ (-G \leq r \leq L-M-1)}} f(\beta[w]_{-G}) = \sum_{\substack{c_r \in \mathbb{F}_q \\ (-G \leq r \leq L-M-1)}} f(\lfloor \beta w \rfloor_{-G+M}) + o(\widehat{G}). \quad (\text{A.9})$$

To establish (A.7), it thus suffices to show that, when L and β are as in the lemma, and G is sufficiently large in terms of L and M , we have

$$\begin{aligned} & \sum_{\substack{a_k \in \mathbb{F}_q \\ (-G+M \leq k \leq L-1)}} f(a_{-G+M} t^{-G+M} + \dots + a_{L-1} t^{L-1}) \\ &= \sum_{\substack{c_r \in \mathbb{F}_q \\ (-G \leq r \leq L-M-1)}} f(\lfloor \beta w \rfloor_{-G+M}) + o(\widehat{G}). \end{aligned} \quad (\text{A.10})$$

To prove (A.10), we compare the coefficients of like powers of t between the expansions for $\lfloor \beta w \rfloor_{-G+M}$ and $a_{-G+M} t^{-G+M} + \dots + a_{L-1} t^{L-1}$. Since $\langle \beta \rangle = \widehat{M}$, we can express β as

$$\beta = \sum_{j=-\infty}^M b_j t^j,$$

where $b_j \in \mathbb{F}_q$ ($j \leq M$) with $b_M \neq 0$. On recalling (A.6), a modicum of computation reveals that

$$\begin{aligned}\Theta_{L-1}(\beta w) &= b_M c_{L-M-1}, \\ \Theta_{L-2}(\beta w) &= b_M c_{L-M-2} + b_{M-1} c_{L-M-1}, \\ &\vdots \\ \Theta_{-G+M}(\beta w) &= b_M c_{-G} + b_{M-1} c_{-G+1} + \dots + b_{-G+2M-L+1} c_{L-M-1}.\end{aligned}$$

Since $b_M \neq 0$, the change of variables from a_k ($-G + M \leq k \leq L - 1$) to c_r ($-G \leq r \leq L - M - 1$) via the linear equations $a_k = \Theta_k(\beta w)$ ($-G + M \leq k \leq L - 1$) is invertible. In other words, as the a_k run through all elements of \mathbb{F}_q , so do the c_r . The relation (A.10) is thus an equality without the error term. Putting (A.10) into (A.9) yields (A.7). This completes the proof of the lemma. \square

With this lemma, we are now ready to settle Theorem A.1, in the case where M is a diagonal matrix.

Lemma A.3. *Theorem A.1 holds for any invertible diagonal matrix M .*

Proof. Suppose the lemma holds for some positive integer D . Now let Γ' be the box given by (A.1), and let Γ be the same box with D replaced by $D + 1$. Let $f : \Gamma \rightarrow \mathbb{C}$ be continuous, and $M \in GL_{D+1}(\mathbb{K}_\infty)$. We express M as

$$M = \begin{pmatrix} \beta_1 & 0 & \cdots & 0 & 0 \\ 0 & \beta_2 & \cdots & 0 & 0 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \beta_D & 0 \\ 0 & 0 & \cdots & 0 & \beta_{D+1} \end{pmatrix},$$

for some $\beta_1, \dots, \beta_{D+1} \in \mathbb{K}_\infty \setminus \{0\}$. Write M' for the $D \times D$ -matrix obtained from the top left corner of M . Then with $\alpha' = (\alpha_1, \dots, \alpha_D)$ and $\alpha = (\alpha_1, \dots, \alpha_{D+1})$, we have

$$\int_{\Gamma} f(\alpha) d\alpha = \int_{\Gamma'} \left(\int_{\langle \alpha_{D+1} \rangle < \widehat{R}_{D+1}} f(\alpha', \alpha_{D+1}) d\alpha_{D+1} \right) d\alpha'.$$

Note that for any fixed $\alpha' \in \Gamma'$, the function $\alpha_{D+1} \mapsto f(\alpha', \alpha_{D+1})$ is continuous on $\mathcal{D}_{R_{D+1}}$. We can thus apply Lemma A.2 to perform the substitution $\alpha_{D+1} = \beta_{D+1} \gamma_{D+1}$ in the α_{D+1} -integral. This, together with an application of Fubini's theorem, yields

$$\begin{aligned}& \int_{\Gamma} f(\alpha) d\alpha \\ &= \langle \beta_{D+1} \rangle \int_{\Gamma'} \int_{\langle \gamma_{D+1} \rangle < \widehat{R}_{D+1} \langle \beta_{D+1} \rangle^{-1}} f(\alpha', \beta_{D+1} \gamma_{D+1}) d\gamma_{D+1} d\alpha' \\ &= \langle \beta_{D+1} \rangle \int_{\langle \gamma_{D+1} \rangle < \widehat{R}_{D+1} \langle \beta_{D+1} \rangle^{-1}} \left(\int_{\Gamma'} f(\alpha', \beta_{D+1} \gamma_{D+1}) d\alpha' \right) d\gamma_{D+1}.\end{aligned}$$

As before, when γ_{D+1} is fixed, the function $\alpha' \mapsto f(\alpha', \beta_{D+1}\gamma_{D+1})$ is continuous on Γ' . By the induction hypothesis, we can therefore apply the change of variables $\alpha_i = \beta_i\gamma_i$ ($i = 1, \dots, D$) to the α' -integral above. With $\gamma = (\gamma_1, \dots, \gamma_{D+1})$ and $\gamma' = (\gamma_1, \dots, \gamma_D)$, this gives

$$\begin{aligned} & \int_{\Gamma} f(\alpha) \, d\alpha \\ &= \left(\prod_{i=1}^{D+1} \langle \beta_i \rangle \right) \int_{\langle \gamma_{D+1} \rangle < \widehat{R}_{D+1} \langle \beta_{D+1} \rangle^{-1}} \int_{M'\gamma' \in \Gamma'} f(\beta_1\gamma_1, \dots, \beta_{D+1}\gamma_{D+1}) \, d\gamma' d\gamma_{D+1}. \end{aligned}$$

On recalling the definitions of the boxes Γ and Γ' , together with those of the matrices M and M' , we simplify this further as

$$\int_{\Gamma} f(\alpha) \, d\alpha = \langle \det M \rangle \int_{M\gamma \in \Gamma} f(M\gamma) \, d\gamma,$$

as required. By induction, Theorem A.1 thus holds for every positive integer D . \square

It therefore remains to settle Theorem A.1 in the case where M is a unipotent matrix. This is accomplished in the following lemma.

Lemma A.4. *Theorem A.1 holds for any unipotent matrix M with entries in \mathbb{K}_{∞} .*

Proof. We prove this by induction on D . The case where $D = 1$ is trivial. Suppose the lemma is true for some positive integer D . Let Γ' and Γ be as in the proof of the previous lemma. Now let $f : \Gamma \rightarrow \mathbb{C}$ be continuous. Take $M \in GL_{D+1}(\mathbb{K}_{\infty})$. Since Theorem A.1 holds for permutation matrices, we can multiply M by permutation matrices without altering either $\langle \det M \rangle$ or the value of the integral on the right side of (A.2). We can thus assume that M is in the form

$$M = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ m_{2,1} & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ m_{D,1} & m_{D,2} & \cdots & 1 & 0 \\ m_{D+1,1} & m_{D+1,2} & \cdots & m_{D+1,D} & 1 \end{pmatrix}$$

where $m_{i,j} \in \mathbb{K}_{\infty}$ for $1 \leq j < i \leq D+1$. Let M' be the $D \times D$ -matrix formed from the entries in the top left hand corner of M . Take any $\gamma \in \mathbb{K}_{\infty}^{D+1}$, and write $\gamma' = (\gamma_1, \dots, \gamma_D)$. An elementary computation reveals that the first D coordinates of the $(D+1)$ -dimensional vector $M\gamma$ form the vector $M'\gamma'$. Crucially, this D -dimensional vector is independent of γ_{D+1} . Meanwhile, the last coordinate of $M\gamma$ is

$$(M\gamma)_{D+1} = \zeta + \gamma_{D+1},$$

where $\zeta = \zeta(\gamma_1, \dots, \gamma_D) = m_{D+1,1}\gamma_1 + \dots + m_{D+1,D}\gamma_D$. Hence we can rewrite the integral on the right side of (A.2) as

$$\int_{M\gamma \in \Gamma} f(M\gamma) \, d\gamma = \int_{M'\gamma' \in \Gamma'} \int_{\langle \zeta + \gamma_{D+1} \rangle < \widehat{R}_{D+1}} f(M'\gamma', \zeta + \gamma_{D+1}) \, d\gamma_{D+1} d\gamma'.$$

Putting $\alpha_{D+1} = \gamma_{D+1} + \zeta$ in the innermost integral, and using Fubini's theorem, we can simplify the above as

$$\begin{aligned} \int_{M\gamma \in \Gamma} f(M\gamma) \, d\gamma &= \int_{M'\gamma' \in \Gamma'} \int_{\langle \alpha_{D+1} \rangle < \widehat{R}_{D+1}} f(M'\gamma', \alpha_{D+1}) \, d\alpha_{D+1} d\gamma' \\ &= \int_{\langle \alpha_{D+1} \rangle < \widehat{R}_{D+1}} \int_{M'\gamma' \in \Gamma'} f(M'\gamma', \alpha_{D+1}) \, d\gamma' d\alpha_{D+1}. \end{aligned}$$

Put $\alpha' = (\alpha_1, \dots, \alpha_D)$. As before, we see that for fixed α_{D+1} , the function $\alpha' \mapsto f(\alpha', \alpha_{D+1}) : \Gamma' \rightarrow \mathbb{C}$ is continuous. Moreover, from the way in which the matrix M' is chosen, we see that M' is a $D \times D$ -unipotent matrix. Using the induction hypothesis on the inner integral above, we have

$$\int_{M\gamma \in \Gamma} f(M\gamma) \, d\gamma = \langle \det M \rangle^{-1} \int_{\langle \alpha_{D+1} \rangle < \widehat{R}_{D+1}} \int_{\Gamma'} f(\alpha', \alpha_{D+1}) \, d\alpha' d\alpha_{D+1}.$$

On recalling the definitions of the boxes Γ and Γ' , and noting that $\det M' = \det M$, we obtain

$$\langle \det M \rangle \int_{M\gamma \in \Gamma} f(M\gamma) \, d\gamma = \int_{\Gamma} f(\alpha) \, d\alpha,$$

as required. This completes the proof of the inductive step, and hence that of the lemma. \square

Theorem A.1, in its full generality, now follows on combining Lemmata A.3 and A.4 together with the trivial case where M is a permutation matrix.

A couple of remarks are in order. First, in the proof of Theorems 1.1 and 1.2, we only use Theorem A.1 in the simple case where f is defined in terms of the exponential given in (2.6), and $M = \beta I_D$, for various $\beta \in \mathbb{K}_\infty \setminus \{0\}$ and positive integers D . Second, we have so far only proved the change-of-variable property when the substitutions are linear. It is natural to ask whether non-linear substitutions are permissible. Unfortunately we do not have this luxury in function fields as we do over \mathbb{R} . As an illustration, suppose we wish to make the substitution $\alpha = \gamma^N$, for some positive integer N . Then $\text{ord } \alpha$ must be divisible by N . So those α with $\text{ord } \alpha$ not divisible by N certainly cannot be substituted in this manner. However, for the purposes of this paper, it suffices to content ourselves with linear substitutions only.

APPENDIX B. GEOMETRY OF NUMBERS OVER FUNCTION FIELDS

In this section, we establish function field analogues of well-known results from the geometry of numbers. It is worth mentioning that Mahler [12] has proved the analogues of Minkowski's theorems over any field with a non-Archimedean valuation. But we deem it worthwhile to prove these results over the field \mathbb{K}_∞ in particular, so as to tailor them for the specific purposes of this paper. The approach taken here is adapted from that of Chapter 12 of [4].

We define a *lattice* in \mathbb{K}_∞^D to be the set of points in the form $\mathbf{x} = \Lambda \mathbf{u}$, where Λ is a $D \times D$ -matrix over \mathbb{K}_∞ , and \mathbf{u} is an integral D -tuple. By a slight abuse of

notation, we also denote the set of such \mathbf{x} by Λ . It is apparent that every lattice is an additive subgroup of \mathbb{K}_∞^D , and has an *integral basis* $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(D)}$, given by the columns of the matrix Λ . Under this basis, every point in this lattice can be expressed as a linear combination of the $\mathbf{x}^{(\nu)}$ ($1 \leq \nu \leq D$) over \mathbb{A} . We define the *determinant* $d(\Lambda)$ of the lattice Λ to be $\langle \det \Lambda \rangle$. Here the Λ in the latter notation refers to the underlying matrix of the lattice. It is easily seen that the determinant of a lattice is independent of the choice of integral basis. Indeed, let $\{\mathbf{v}^{(\nu)} : \nu = 1, \dots, D\}$ and $\{\mathbf{w}^{(\nu)} : \nu = 1, \dots, D\}$ be two integral bases of Λ . Write $V = (\mathbf{v}^{(1)}, \dots, \mathbf{v}^{(D)})$ and $W = (\mathbf{w}^{(1)}, \dots, \mathbf{w}^{(D)})$. We can then find matrices $Z, Z' \in GL_D(\mathbb{A})$ such that $V = ZW$ and $W = Z'V$. Thus $V = ZZ'V$, which in turn yields $ZZ' = I_D$. It follows that $\det Z$ is a unit in \mathbb{A} , which has absolute value 1 in \mathbb{K}_∞ . This shows that the determinant of a lattice as defined above is independent of the choice of integral basis. Meanwhile, it follows from the definition (1.3) of the absolute value on \mathbb{K}_∞^D that every D -dimensional box, in the sense of Appendix A, is an additive subgroup of \mathbb{K}_∞^D .

We first need a lemma concerning the volume of any parallelepiped in the shape

$$\mathcal{P}(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(D)}) = \{\alpha_1 \mathbf{x}^{(1)} + \dots + \alpha_D \mathbf{x}^{(D)} : \alpha_1, \dots, \alpha_D \in \mathbb{T}\},$$

for any linearly independent vectors $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(D)} \in \mathbb{K}_\infty^D$.

Lemma B.1. *Let D be any positive integer. Let $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(D)}$ be real D -tuples that are linearly independent over \mathbb{K}_∞ . Then*

$$\text{Vol } \mathcal{P}(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(D)}) = \langle \det(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(D)}) \rangle.$$

Proof. The claim here is a straightforward consequence of Theorem A.1. Write M for the $D \times D$ -matrix with columns $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(D)}$. Since $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(D)}$ are linearly independent, it follows that $M \in GL_D(\mathbb{K}_\infty)$. So every element of $\mathcal{P} = \mathcal{P}(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(D)})$ is in the form $\boldsymbol{\beta} = M\boldsymbol{\alpha}$, where $\boldsymbol{\alpha} \in \mathbb{T}^D$. Using Theorem A.1, we can apply this substitution in the integral $\int_{\mathbb{T}^D} d\boldsymbol{\alpha}$, which equals 1 by normalisation. Hence

$$1 = \int_{\mathbb{T}^D} d\boldsymbol{\alpha} = \langle \det M \rangle^{-1} \int_{\mathcal{P}} d\boldsymbol{\beta} = \langle \det M \rangle^{-1} \text{Vol } \mathcal{P}.$$

The required equality thus follows by rearranging this equality, and recalling the definition of M . \square

In the remainder of this appendix, we fix a lattice Λ in \mathbb{K}_∞^D . Using this lemma, we derive the following analogue of Minkowski's first theorem over function fields.

Lemma B.2. *Suppose Γ is an D -dimensional box with*

$$\text{Vol } \Gamma > d(\Lambda). \tag{B.1}$$

Then Γ contains a non-zero point of Λ .

Proof. Let $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(D)}$ denote an integral basis for Λ . Put $\mathcal{P} = \mathcal{P}(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(D)})$. For each $\mathbf{x} \in \Lambda$, let $\mathcal{R}(\mathbf{x}) = \mathcal{P} \cap (\Gamma - \mathbf{x})$. By the translation-invariance of the measure on \mathbb{K}_∞^D , we have

$$\text{Vol } \mathcal{R}(\mathbf{x}) = \text{Vol } ((\mathcal{P} + \mathbf{x}) \cap \Gamma)$$

for each such \mathbf{x} . Summing over all such \mathbf{x} gives

$$\sum_{\mathbf{x} \in \Lambda} \text{Vol } \mathcal{R}(\mathbf{x}) = \sum_{\mathbf{x} \in \Lambda} \int_{\mathcal{P} + \mathbf{x}} 1_\Gamma(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha}.$$

Here 1_Γ denotes the indicator function on Γ . The compactness of Γ implies that, on both sides of the last equation, only finitely many non-zero terms occur, so there is no issue with the convergence of the two infinite sums involved. Also, from the definition of \mathcal{P} , it transpires that the translates $\mathcal{P} + \mathbf{x}$ of \mathcal{P} by lattice points $\mathbf{x} \in \Lambda$ are disjoint, and their union coincides with \mathbb{K}_∞^D . It thus follows from the last displayed equality, together with (B.1) and Lemma B.1, that

$$\sum_{\mathbf{x} \in \Lambda} \text{Vol } \mathcal{R}(\mathbf{x}) = \int_{\mathbb{K}_\infty^s} 1_\Gamma(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} = \text{Vol } \Gamma > d(\Lambda) = \text{Vol } \mathcal{P}.$$

Since each $\mathcal{R}(\mathbf{x})$ is a subset of \mathcal{P} , it follows from this inequality that two of the $\mathcal{R}(\mathbf{x})$ must intersect. There thus exist distinct lattice points $\mathbf{x}_1, \mathbf{x}_2 \in \Lambda$ and $\boldsymbol{\alpha} \in \mathcal{P}$ for which $\boldsymbol{\alpha} \in \mathcal{R}(\mathbf{x}_1) \cap \mathcal{R}(\mathbf{x}_2)$. The definition of $\mathcal{R}(\mathbf{x})$ gives $\boldsymbol{\alpha} + \mathbf{x}_1 \in \Gamma$ and $\boldsymbol{\alpha} + \mathbf{x}_2 \in \Gamma$. Since Γ is an additive group, we must have

$$\mathbf{x}_1 - \mathbf{x}_2 = (\boldsymbol{\alpha} + \mathbf{x}_1) - (\boldsymbol{\alpha} + \mathbf{x}_2) \in \Gamma.$$

This point must be non-zero since $\mathbf{x}_1 \neq \mathbf{x}_2$. Since $\mathbf{x}_1, \mathbf{x}_2 \in \Lambda$, we also have $\mathbf{x}_1 - \mathbf{x}_2 \in \Lambda$, as required. \square

Before advancing further, we need to introduce some notation and terminology. When μ is a positive integer, and $\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(\mu)} \in \mathbb{K}_\infty^D$, let $\langle \mathbf{y}^{(1)}, \dots, \mathbf{y}^{(\mu)} \rangle_{\mathbb{K}}$ denote the set of all linear combinations of $\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(\mu)}$ over \mathbb{K} . Meanwhile, define \widehat{R}_1 to be the infimum of $\langle \mathbf{x} \rangle$ over all non-zero points $\mathbf{x} \in \Lambda$. Take $\mathbf{x}^{(1)}$ to be some point in Λ with $\langle \mathbf{x}^{(1)} \rangle = \widehat{R}_1$. Inductively, for every $\nu = 2, \dots, D$, define \widehat{R}_ν to be the infimum of $\langle \mathbf{x} \rangle$ over all points $\mathbf{x} \in \Lambda \setminus \langle \mathbf{x}^{(1)}, \dots, \mathbf{x}^{(\nu-1)} \rangle_{\mathbb{K}}$. Take $\mathbf{x}^{(\nu)} \in \Lambda$ with $\langle \mathbf{x}^{(\nu)} \rangle = \widehat{R}_\nu$. We call \widehat{R}_ν the ν -th *successive minima* of Λ , and $\mathbf{x}^{(\nu)}$ a ν -th *minimal point* of Λ . Evidently $\widehat{R}_1 \leq \dots \leq \widehat{R}_D$.

Lemma B.3. *Let $\widehat{R}_1, \dots, \widehat{R}_D$ be the successive minima of Λ . Then*

$$\widehat{R}_1 \dots \widehat{R}_D = d(\Lambda).$$

Proof. By applying a suitable unimodular linear transformation if necessary, we can assume without loss of generality that Λ has minimal points in the

form

$$\begin{aligned}\mathbf{x}^{(1)} &= (x_1^{(1)}, 0, \dots, 0), \\ \mathbf{x}^{(2)} &= (x_1^{(2)}, x_2^{(2)}, \dots, 0) \\ &\vdots \\ \mathbf{x}^{(D)} &= (x_1^{(D)}, x_2^{(D)}, \dots, x_D^{(D)}),\end{aligned}$$

for some $x_\mu^{(\nu)} \in \mathbb{K}_\infty$ ($1 \leq \mu \leq \nu \leq D$). Choose an arbitrary integral basis $\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(D)}$ of Λ . Then we can find a matrix $Z \in GL_D(\mathbb{A})$ such that

$$(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(D)}) = Z(\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(D)}).$$

Taking absolute values of the determinants of both sides above, and recalling that $d(\Lambda) = \langle \det(\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(D)}) \rangle$, we get

$$\langle \det(\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(D)}) \rangle = \langle \det Z \rangle d(\Lambda).$$

Since $\det Z$ is a non-zero element of \mathbb{A} , it follows that $\langle \det Z \rangle \geq 1$. This together with our initial assumption on the points $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(D)}$ gives

$$d(\Lambda) \leq \langle x_1^{(1)} \dots x_D^{(D)} \rangle.$$

Courtesy of (1.3), the above implies that

$$d(\Lambda) \leq \langle \mathbf{x}^{(1)} \rangle \dots \langle \mathbf{x}^{(D)} \rangle.$$

From the definition of minimal points, this gives

$$d(\Lambda) \leq \widehat{R}_1 \dots \widehat{R}_D.$$

It thus remains to establish the reverse inequality.

To this end, let Γ be the box given by the inequalities (A.1). We claim that Γ contains no non-zero point of Λ . Indeed, suppose Γ contains a non-zero point \mathbf{x} of Λ . From the definition of minimal points, we can find a $\nu = 1, \dots, D$ such that \mathbf{x} is linearly dependent over \mathbb{K} on $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(\nu)}$, but not on $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(\nu-1)}$. So on one hand, it follows from the definition of \widehat{R}_ν that $\langle \mathbf{x} \rangle \geq \widehat{R}_\nu$. On the other hand, by the way in which we have chosen the minimal points $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(\nu)}$, we deduce that $x_{\nu+1} = \dots = x_D = 0$. Also, since $\mathbf{x} \in \Gamma$, we have $\langle x_u \rangle < \widehat{R}_u$ for $u = 1, \dots, \nu$. On recalling that $\widehat{R}_1 \leq \dots \leq \widehat{R}_\nu$, we see that $\langle \mathbf{x} \rangle < \widehat{R}_\nu$. This is a contradiction to what we mentioned above. It thus follows from (A.1) that $\mathbf{x} \notin \Gamma$, which proves the claim. This claim, together with Lemma B.2, implies that $\text{Vol } \Gamma \leq d(\Lambda)$. But it follows easily from (2.9) that $\text{Vol } \Gamma = \widehat{R}_1 \dots \widehat{R}_D$, so we have the desired inequality

$$d(\Lambda) \geq \widehat{R}_1 \dots \widehat{R}_D.$$

□

Although the choice of minimal points for a lattice is not unique, the following variant of Lemma 12.3 of [4] gives an essentially canonical choice that is convenient to work with.

Lemma B.4. *Let Λ and $\widehat{R}_1, \dots, \widehat{R}_D$ be as in Lemma B.3. Then Λ has an integral basis $\mathbf{X}^{(\nu)}$ ($1 \leq \nu \leq D$) in the form*

$$\mathbf{X}^{(\nu)} = (X_1^{(\nu)}, \dots, X_\nu^{(\nu)}, 0, \dots, 0) \quad (\nu = 1, \dots, D) \quad (\text{B.2})$$

such that for each such ν , we have

$$\langle \mathbf{X}^{(\nu)} \rangle = \widehat{R}_\nu \quad \text{and} \quad \langle X_\nu^{(\nu)} \rangle = \widehat{R}_\nu. \quad (\text{B.3})$$

Proof. Let $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(D)}$ be chosen as in the proof of Lemma B.3. Take $\mathbf{X}^{(1)} = \mathbf{x}^{(1)}$, and choose $\mathbf{X}^{(2)} \in \Lambda \cap \langle \mathbf{x}^{(1)}, \mathbf{x}^{(2)} \rangle_{\mathbb{K}}$, such that $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ generate integrally all the points in $\Lambda \cap \langle \mathbf{x}^{(1)}, \mathbf{x}^{(2)} \rangle_{\mathbb{K}}$. The choice of $\mathbf{X}^{(2)}$ is arbitrary to the extent of added integral multiples of $\mathbf{x}^{(1)}$. Since $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ generate rationally (not necessarily integrally) all the points of Λ , in particular we can find $m, u_1, u_2 \in \mathbb{A}$ with $m \neq 0$ such that

$$m\mathbf{X}^{(2)} = u_1\mathbf{x}^{(1)} + u_2\mathbf{x}^{(2)}. \quad (\text{B.4})$$

Our aim now is to show that the polynomials u_1 and u_2 above can be assumed to have absolute values at most $\langle m \rangle$. This then implies that

$$\begin{aligned} \langle \mathbf{X}^{(2)} \rangle &\leq \max \{ \langle u_1/m \rangle \langle \mathbf{x}^{(1)} \rangle, \langle u_2/m \rangle \langle \mathbf{x}^{(2)} \rangle \} \\ &\leq \max \{ \widehat{R}_1, \widehat{R}_2 \} = \widehat{R}_2. \end{aligned}$$

To prove our claim, note that $\mathbf{x}^{(2)}$ is generated integrally by $\mathbf{X}^{(1)} = \mathbf{x}^{(1)}$ and $\mathbf{X}^{(2)}$, so we can find $v_1, v_2 \in \mathbb{A}$ such that

$$\mathbf{x}^{(2)} = v_1\mathbf{x}^{(1)} + v_2\mathbf{X}^{(2)}.$$

Combining this with (B.4) gives

$$(m - u_2v_2)\mathbf{X}^{(2)} = (u_1 + u_2v_1)\mathbf{x}^{(1)}. \quad (\text{B.5})$$

Due to our choice of $\mathbf{x}^{(1)}$ in Lemma B.3, by comparing the second coordinates on both sides above, we get

$$m = u_2v_2. \quad (\text{B.6})$$

Meanwhile, on considering the first coordinates in (B.5), we see that $u_1 = -u_2v_1$. Putting this and (B.6) into (B.4) yields

$$u_2v_2\mathbf{X}^{(2)} = u_2(-v_1\mathbf{x}^{(1)} + \mathbf{x}^{(2)}). \quad (\text{B.7})$$

If $u_2 = 0$, then by looking at the second coordinates on both sides of (B.4), we get $m = 0$, a contradiction. Hence $u_2 \neq 0$, and (B.7) gives

$$v_2\mathbf{X}^{(2)} = -v_1\mathbf{x}^{(1)} + \mathbf{x}^{(2)}. \quad (\text{B.8})$$

Arguing as before, we have $v_2 \neq 0$. Write $-v_1 = w_1v_2 + r_1$, where $w_1, r_1 \in \mathbb{A}$ with $\langle r_1 \rangle < \langle v_2 \rangle$. Using this in (B.8), we have

$$v_2\mathbf{X}^{(2)} = r_1\mathbf{x}^{(1)} + \mathbf{x}^{(2)}. \quad (\text{B.9})$$

Here we have redefined $\mathbf{X}^{(2)}$ by subtracting $w_1\mathbf{x}^{(1)}$ from the old $\mathbf{X}^{(2)}$. On comparing (B.9) with (B.4), our claim in this paragraph is thus verified.

In the same vein, for each $\nu = 3, \dots, D$, we define $\mathbf{X}^{(\nu)}$ to be a point in $\Lambda \cap \langle \mathbf{x}^{(1)}, \dots, \mathbf{x}^{(\nu)} \rangle_{\mathbb{K}}$ such that $\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(\nu)}$ generate integrally all points in $\Lambda \cap \langle \mathbf{x}^{(1)}, \dots, \mathbf{x}^{(\nu)} \rangle_{\mathbb{K}}$, with $\mathbf{X}^{(\nu)}$ arbitrary with respect to added integral multiples of $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(\nu-1)}$. A similar argument to the one used in the last paragraph shows that

$$\langle \mathbf{X}^{(\nu)} \rangle \leq \widehat{R}_\nu \quad \text{for all } \nu = 1, \dots, D,$$

and consequently

$$\langle X_\nu^{(\nu)} \rangle \leq \langle \mathbf{X}^{(\nu)} \rangle \leq \widehat{R}_\nu$$

for all such ν . Our choice of the $\mathbf{X}^{(\nu)}$ ($1 \leq \nu \leq s$) indicates that these points are in the form given in (B.3).

The lower bound $\langle X_\nu^{(\nu)} \rangle \geq \widehat{R}_\nu$ follows by comparing determinants. Our choice of $\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(D)}$ above gives rise to an integral basis of Λ , so

$$d(\Lambda) = \langle \det(\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(D)}) \rangle = \langle X_1^{(1)} \dots X_D^{(D)} \rangle.$$

Recalling the definition of successive minima, we have

$$d(\Lambda) \leq \widehat{R}_1 \dots \widehat{R}_{\nu-1} \langle X_\nu^{(\nu)} \rangle \widehat{R}_{\nu+1} \dots \widehat{R}_D.$$

Applying Lemma B.3, we obtain

$$\langle X_\nu^{(\nu)} \rangle \geq \widehat{R}_\nu \quad \text{for all } \nu = 1, \dots, D.$$

Hence $\langle \mathbf{X}^{(\nu)} \rangle \geq \langle X_\nu^{(\nu)} \rangle \geq \widehat{R}_\nu$, as required. \square

The following lemma counts exactly the number of lattice points lying within a given distance from the origin.

Lemma B.5. *Let Λ and $\widehat{R}_1, \dots, \widehat{R}_D$ be as in Lemma B.3. Then*

$$\# \left\{ \mathbf{x} \in \Lambda : \langle \mathbf{x} \rangle < \widehat{U} \right\} = \begin{cases} 1, & \text{if } U < R_1, \\ \widehat{U}^\nu (\widehat{R}_1 \dots \widehat{R}_\nu)^{-1}, & \text{if } R_\nu \leq U < R_{\nu+1}. \end{cases}$$

Proof. When $U < R_1$, the only point in Λ with absolute value less than \widehat{U} is 0, so the lemma follows immediately in this case. Now suppose $R_\nu \leq U < R_{\nu+1}$. By Lemma B.4, every $\mathbf{x} \in \Lambda$ with $\langle \mathbf{x} \rangle < \widehat{U}$ has the form

$$\mathbf{x} = v_1 \mathbf{X}^{(1)} + \dots + v_\nu \mathbf{X}^{(\nu)}, \tag{B.10}$$

where $v_1, \dots, v_\nu \in \mathbb{A}$. The normalisation (B.2) in that lemma implies further that

$$\begin{aligned} \langle x_1 \rangle &= \langle v_1 X_1^{(1)} + v_2 X_1^{(2)} + \dots + v_\nu X_1^{(\nu)} \rangle < \widehat{U}, \\ \langle x_2 \rangle &= \langle v_2 X_2^{(2)} + \dots + v_\nu X_2^{(\nu)} \rangle < \widehat{U}, \\ &\vdots \\ \langle x_\nu \rangle &= \langle v_\nu X_\nu^{(\nu)} \rangle < \widehat{U}. \end{aligned} \tag{B.11}$$

The last inequality above together with (B.3) yields

$$\langle v_\nu \rangle < \widehat{U} \widehat{R}_\nu^{-1}.$$

Inductively, for any $\sigma = 2, \dots, \nu$, once we obtain

$$\langle v_\mu \rangle < \widehat{U} \widehat{R}_\mu^{-1}$$

for all $\mu = \sigma, \dots, \nu$, we know from Lemma B.4 that

$$\langle v_\mu X_{\sigma-1}^{(\mu)} \rangle < \widehat{U} \widehat{R}_\mu^{-1} \langle \mathbf{X}^{(\mu)} \rangle = \widehat{U}$$

for all such μ . The $(\sigma - 1)$ -th inequality in (B.11), namely

$$\langle x_{\sigma-1} \rangle = \langle v_{\sigma-1} X_{\sigma-1}^{(\sigma-1)} + v_\sigma X_{\sigma-1}^{(\sigma)} + \dots + v_\nu X_{\sigma-1}^{(\nu)} \rangle < \widehat{U},$$

then produces the upper bound

$$\langle v_{\sigma-1} X_{\sigma-1}^{(\sigma-1)} \rangle < \widehat{U}.$$

Applying Lemma B.4 again gives

$$\langle v_{\sigma-1} \rangle < \widehat{U} \widehat{R}_{\sigma-1}^{-1}.$$

Hence v_1, \dots, v_ν are subject to the constraints

$$\langle v_\mu \rangle < \widehat{U} \widehat{R}_\mu^{-1} \quad \text{for } \mu = 1, \dots, \nu. \quad (\text{B.12})$$

Therefore

$$\# \left\{ \mathbf{x} \in \Lambda : \langle \mathbf{x} \rangle < \widehat{U} \right\} \leq \prod_{\mu=1}^{\nu} (\widehat{U} / \widehat{R}_\mu) = \widehat{U}^\nu / (\widehat{R}_1 \dots \widehat{R}_\nu).$$

The reverse inequality can be obtained by counting only those points \mathbf{x} in the form (B.10) which satisfy (B.12). \square

Before we proceed further, we introduce the notion of adjoint lattices. We say that two lattices Λ and M in \mathbb{K}_∞^D are adjoint if their underlying matrices satisfy

$$\Lambda^T M = I_D, \quad (\text{B.13})$$

where Λ^T denotes the transpose of the matrix Λ . The points on Λ and M have the respective forms $\mathbf{x} = \Lambda \mathbf{u}$ and $\mathbf{y} = M \mathbf{v}$, where $\mathbf{u}, \mathbf{v} \in \mathbb{A}^D$. Equation (B.13) then implies that

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{u} \cdot \mathbf{v} \in \mathbb{A}. \quad (\text{B.14})$$

The following lemma gives an inverse relation between the successive minima of two adjoint lattices.

Lemma B.6. *Let Λ and M be adjoint lattices in \mathbb{K}_∞^D , with successive minima $\widehat{R}_1, \dots, \widehat{R}_D$ and $\widehat{S}_1, \dots, \widehat{S}_D$. Then for each $\nu = 1, \dots, D$, we have*

$$\widehat{R}_\nu = \widehat{S}_{D-\nu+1}^{-1}.$$

Proof. Let $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(D)}$ and $\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(D)}$ denote respectively minimal points of Λ and M , with $\langle \mathbf{x}^{(\mu)} \rangle = \widehat{R}_\mu$ and $\langle \mathbf{y}^{(\mu)} \rangle = \widehat{S}_\mu$ for any $\mu = 1, \dots, D$. Fix $\nu = 1, \dots, D$, and let

$$\mathcal{U} = \left\{ \mathbf{y} \in M : \mathbf{x}^{(\mu)} \cdot \mathbf{y} = 0 \text{ for all } \mu = 1, \dots, \nu \right\}.$$

This is an $(D - \nu)$ -dimensional subspace of \mathbb{K}_∞^D . So not all of $\mathbf{y}^{(1)}, \dots, \mathbf{y}^{(D-\nu+1)}$ lie in \mathcal{U} . There thus exist positive integers m and n with $m \leq \nu$ and $n \leq D - \nu + 1$ such that $\mathbf{x}^{(m)} \cdot \mathbf{y}^{(n)} \neq 0$. Owing to (B.14), this implies that

$$\langle \mathbf{x}^{(m)} \cdot \mathbf{y}^{(n)} \rangle \geq 1.$$

But for each $\mu = 1, \dots, D$ we have

$$\langle x_\mu^{(m)} \rangle \leq \langle \mathbf{x}^{(m)} \rangle = \widehat{R}_m \leq \widehat{R}_\nu$$

and

$$\langle y_\mu^{(n)} \rangle \leq \langle \mathbf{y}^{(n)} \rangle = \widehat{S}_n \leq \widehat{S}_{D-\nu+1}.$$

Therefore

$$\widehat{R}_\nu \widehat{S}_{D-\nu+1} \geq 1 \quad \text{for all } \nu = 1, \dots, D. \quad (\text{B.15})$$

The reverse inequality can be proved as follows. From (B.13), we have

$$d(\Lambda)d(M) = 1.$$

Applying Lemma B.3 on the lattices Λ and M in turn gives

$$(\widehat{R}_1 \dots \widehat{R}_D)(\widehat{S}_1 \dots \widehat{S}_D) = 1.$$

Rearranging this yields

$$\widehat{R}_\nu \widehat{S}_{D-\nu+1} = \prod_{\substack{\mu=1 \\ \mu \neq \nu}}^D (\widehat{R}_\mu \widehat{S}_{D-\mu+1})^{-1}.$$

Applying (B.15) leads to

$$\widehat{R}_\nu \widehat{S}_{D-\nu+1} \leq 1 \quad \text{for all } \nu = 1, \dots, D,$$

as required. \square

REFERENCES

- [1] A. G. Aleksandrov and B. Z. Moroz, *Complete intersections in relation to a paper of B. J. Birch*, Bull. London Math. Soc. **34** (2002), no. 2, 149154.
- [2] B. J. Birch, *Forms in many variables*, Proc. Roy. Soc. (A) **265** (1961/62), 245-263.
- [3] M. Car, *Sommes d'exponentielles dans $\mathbb{F}_{2^h}((X^{-1}))$* , Acta Arith. **62** (1992), no. 4, 303-328.
- [4] H. Davenport, *Analytic methods for Diophantine equations and Diophantine inequalities*, Cambridge University Press, Cambridge, 2nd edition, 2005.
- [5] M. Greenberg, *Lectures on Forms in Many Variables*, Benjamin, New York (1969).
- [6] R. Hartshorne, *Algebraic Geometry*, Grad. Texts in Mathematics, **52**, Springer-Verlag, New York, 1977.
- [7] B. Hassett and Y. Tschinkel, *Weak approximation for hypersurfaces of low degree*, Algebraic geometry - Seattle 2005. Part 2, 937-955, Proc. Sympos. Pure Math., 80, Part 2, Amer. Math. Soc., Providence, RI, 2009.
- [8] D. R. Heath-Brown, *Cubic forms in 14 variables*, Invent. Math. **170** (2007), 199-230.
- [9] J-I. Igusa, *An Introduction to the Theory of Local Zeta Functions*, AMS/IP Studies in Advanced Mathematics, 14. American Mathematical Society, Providence, RI; International Press, Cambridge, MA, 2000.
- [10] R. M. Kubota, *Waring's Problem for $\mathbb{F}_q[x]$* , Dissertationes Math. (Rozprawy Mat.) **117** (1974), 60 pp.

- [11] Y.-R. Liu and T. D. Wooley, *Waring's Problem in Function Fields*, J. Reine Angew. Math. **638** (2010), 1-67.
- [12] K. Mahler, *An analogue to Minkowski's geometry of numbers in a field of series*, Annals of Math. **42** (1941), no. 2, 488-522.
- [13] M. Rosen, *Number theory in function fields*, Graduate Texts in Mathematics, vol. 210, Springer-Verlag, New York, 2002.
- [14] W. Schmidt, *Simultaneous rational zeros of quadratic forms*, Seminar Delange-Pisot-Poitou 1981. Progress in Math. **22**, 281-307.
- [15] W. Schmidt, *The density of integer points on homogeneous varieties*, Acta Arith. **154** (1985), no. 3-4, 243-296.
- [16] C. M. Skinner, *Forms over number fields and weak approximation*, Composito Mathematica **106** (1997), 11-29.

SCHOOL OF MATHEMATICS, UNIVERSITY OF BRISTOL, UNIVERSITY WALK, BRISTOL
BS8 1TW, UNITED KINGDOM

E-mail address: `sl5701@bris.ac.uk`